IMPROVED WIRELESS SECURITY THROUGH PHYSICAL LAYER PROTOCOL MANIPULATION AND RADIO FREQUENCY FINGERPRINTING

DISSERTATION

Benjamin W. Ramsey, Captain, USAF

AFIT-ENG-DS-14-S-10

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

DISTRIBUTION STATEMENT A:
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

The views expressed in this dissertation are those of the author and do not reflect the official policy or position of the United States Air Force, the Department of Defense, or the United States Government.

This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

AFIT-ENG-DS-14-S-10

IMPROVED WIRELESS SECURITY THROUGH PHYSICAL LAYER PROTOCOL MANIPULATION AND RADIO FREQUENCY FINGERPRINTING

DISSERTATION

Presented to the Faculty
Graduate School of Engineering
and Management
Air Force Institute of Technology
Air University
Air Education and Training Command
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

Benjamin W. Ramsey, BSEE, MS, MSEE
Captain, USAF

September 2014

DISTRIBUTION STATEMENT A:
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

AFIT-ENG-DS-14-S-10

IMPROVED WIRELESS SECURITY THROUGH PHYSICAL LAYER PROTOCOL MANIPULATION AND RADIO FREQUENCY FINGERPRINTING

DISSERTATION

Benjamin W. Ramsey, BSEE, MS, MSEE
Captain, USAF

Approved:

/signed/ 8 Aug 2014
Barry E. Mullins, PhD (Chairman) Date

/signed/ 8 Aug 2014
Michael A. Temple, PhD (Member) Date

/signed/ 8 Aug 2014
Michael R. Grimaila, PhD, CISM, CISSP (Member) Date

Accepted:

/signed/ 12 Aug 2014
ADEDEJI B. BADIRU, PhD Date
Dean, Graduate School of Engineering
and Management

AFIT-ENG-DS-14-S-10

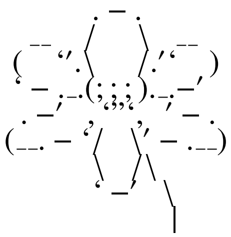
Abstract

Wireless networks are particularly vulnerable to spoofing and route poisoning attacks due to the contested transmission medium. Traditional bit-layer defenses including encryption keys and MAC address control lists are vulnerable to extraction and identity spoofing, respectively. This dissertation explores three novel strategies to leverage the wireless physical layer to improve security in low-rate wireless personal area networks. The first, physical layer protocol manipulation, identifies true transceiver design within remote devices through analysis of replies in response to packets transmitted with modified physical layer headers. Results herein demonstrate a methodology that correctly differentiates among six IEEE 802.15.4 transceiver classes with greater than 99% accuracy, regardless of claimed bit-layer identity. The second strategy, radio frequency fingerprinting, accurately identifies the true source of every wireless transmission in a network, even among devices of the same design and manufacturer. Results suggest that even low-cost signal collection receivers can achieve greater than 90% authentication accuracy within a defense system based on radio frequency fingerprinting. The third strategy, based on received signal strength quantification, can be leveraged to rapidly locate suspicious transmission sources and to perform physical security audits of critical networks. Results herein reduce mean absolute percentage error of a widely-utilized distance estimation model 20% by examining signal strength measurements from real-world networks in a military hospital and a civilian hospital.

To my children

Acknowledgments

I thank my incredible wife for her steadfast love and support.



I also thank my committee for helping me fulfill this childhood dream.

Benjamin W. Ramsey

Table of Contents

List of Figures

Figure Page

Figure Page

List of Tables

IMPROVED WIRELESS SECURITY THROUGH PHYSICAL LAYER PROTOCOL MANIPULATION AND RADIO FREQUENCY FINGERPRINTING

I. Introduction

Modern wireless communications networks have revolutionized the ways in which information is shared. Inexpensive low-rate wireless personal area networks (LR-WPANs) can readily connect thousands of remote devices. Critical infrastructure, from hospitals to smart grids and petroleum refineries, increasingly leverage low-cost wireless connectivity in daily operations. Significant disruptions to these systems could endanger patient lives or cause industrial sabotage, meanwhile open source tools designed to attack and degrade wireless networks have rapidly grown in sophistication and effectiveness. With physical security a low priority in the design of low-cost wireless systems, novel defensive measures are necessary to protect these vulnerable networks.

1.1 Background

Many wireless security architectures rely upon safekeeping of symmetric keys to uphold message confidentiality, message integrity, and device authentication. While the small size and low complexity of LR-WPAN hardware make them effective to deploy in large numbers, these traits also result in tight limitations on device memory and computation. A single network key is shared by every device in the LR-WPAN, although device-to-device confidentiality is also possible using link keys at the application layer. Small, inexpensive wireless nodes lack robust defense against theft and tampering, resulting in physical vulnerabilities to key confidentiality. Key extraction from first and second-generation ZigBee chips is shown to be straightforward [Goo09], and inexpensive tools have recently been developed for locating LR-WPAN devices [RMW12][KP12]. Keys may also be compromised through social engineering, or intercepted (if transmitted to end devices without encryption) by open source tools such as KillerBee [Wri09] and Api-do [GBM^{+}12]. Wireless device identities including Medium Access Control (MAC) or network addresses are readily spoofed by such tools. While bit-layer identities and keys are susceptible to spoofing attacks, physical layer attributes are significantly more difficult for an attacker to mimic.

1.2 Research Goal

The overarching goal of this dissertation is to investigate state-of-the-art methodologies for exploiting wireless physical layer features to improve network security. The hypothesis herein is that physical layer features can be successfully exploited.

1.3 Research Approach

This dissertation investigates novel methods for leveraging the IEEE 802.15.4 physical layer protocol to improve LR-WPAN security and cyberspace situational awareness. To this end three classes of techniques are examined: protocol manipulation, radio frequency fingerprinting, and received signal strength analysis. The three techniques are useful individually or can be combined for wireless defense-in-depth.

1.3.1 Physical Layer Protocol (PHY) Manipulation.

The IEEE 802.15.4 PHY specifies radio frequency and message synchronization standards for LR-WPAN communication. Transceiver manufacturers design their hardware to comply with this standard, but each design differs somewhat in its implementation. Ramsey and Mullins [RM13] first reported that packet reception could be degraded or completely blocked by manipulating the IEEE 802.15.4 PHY header of outdoing transmissions, and that this behavior differed significantly among transceiver designs. This opened a new field of research wherein PHY manipulation is leveraged to protect sensitive data, identify remote transceiver hardware, intrusion detection, and provide PHY-augmented authentication.

1.3.2 Radio Frequency (RF) Fingerprinting.

Transmitter fingerprinting originally targeted military radar stations, and similar techniques have been adapted in recent years to wireless communications networks. Many physical layer features have been used to generated RF fingerprints to uniquely identify transmitters, including frequency, phase, and amplitude metrics. Costs associated with RF fingerprinting have historically been high due to the receiver sensitivity required for precise physical layer measurements. However, low-cost software-defined radios show promise toward making RF fingerprinting more accessible. This dissertation investigates and compares the performance of both high-cost and low-cost RF fingerprinting receivers.

1.3.3 Received Signal Strength Indicator (RSSI).

RSSI is the quantification of RF power input to a wireless receiver. This unitless metric decreases with increased distance from the transmitter. If the transmit power and RSSI are both known, then the distance from receiver to transmitter can be estimated from RSSI alone. The open-source KillerBee framework for exploiting and exploring IEEE 802.15.4 networks includes a tool named `zbfind` that uses a log-distance path loss model to estimate distance to nearby LR-WPAN transmitters. This dissertation investigates the performance of this tool and makes recommendations to improve its accuracy.

1.4 Dissertation Organization

This dissertation has four main chapters, each of which is closely based on research papers that have either been published at academic conferences or that are undergoing peer review for archival journal publication. Chapter 2 explores PHY manipulation for intrusion detection, device fingerprinting, and device authentication applications; the chapter is a significant extension of work presented at the International Federation for Information Processing (IFIP) working group 11.10 International Conference on Critical

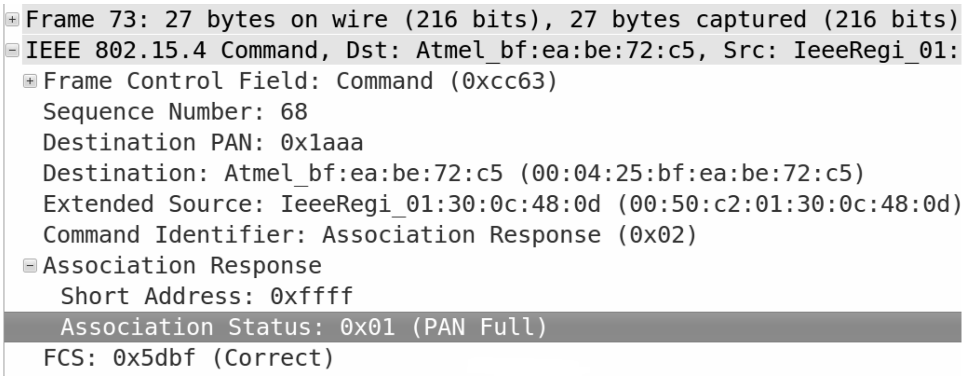
Infrastructure Protection [RM13]. Chapter 3 is an extension of work published at the IEEE Global Communications Conference [RTM12], in which IEEE 802.15.4 transmitters are uniquely identified by RF fingerprints generated from PHY preamble responses. Chapter 4 extends the work in Chapter 3 by performing comparative RF fingerprinting experiments with high-cost and low-cost signal receivers. Chapter 5 presents results toward improving a popular transmitter distance estimation model based on RSSI; the results will be presented at the 2014 Military Communications Conference [RMLS14]. Chapter 6 concludes and proposes future work.

II. Wireless Intrusion Detection and Device Fingerprinting through PHY Manipulation

2.1 Introduction

Properly securing low-rate LR-WPANs is challenging due to tight resource constraints. LR-WPAN hardware is designed to be as inexpensive as possible, and tamper resistance was not an early vendor priority; first and second-generation ZigBee chips were found to be vulnerable to encryption key extraction [Goo09]. Flash memory available for application development is typically limited to less than 100 kB, e.g., 48 kB on the TmoteSky mote [DDT11] and 60 kB on the Freescale MC13213 [Fre08]. With flash at a premium, application developer guides even discourage the use of security: "Do not use a secure network unless required. ZigBee security is about 8K" [Fre08]. Security headers increase packet overhead, expending additional wireless transmission energy and presenting a trade-off for LR-WPANs reliant upon battery power. IEEE 802.15.4 leaves key establishment to higher layers, such as the ZigBee stack, yet the entire LR-WPAN can be compromised if keys are mishandled. Support for access control lists varies substantially among LR-WPAN chipsets as well; the CC2420 only supports two device entries [NW04].

Any network keys wirelessly distributed in plain text to end nodes can be intercepted by eavesdroppers. The open source KillerBee framework for exploiting IEEE 802.15.4 LR-WPANs [Wri09] includes a script (*zbdsniff*) that extracts any observed keys from wireless capture files. KillerBee also includes tools for message replay attacks (`zbreplay`), transmitter tracking (`zbfind`), and denial-of-service attacks (`zbassocflood`).



```
⊞ Frame 73: 27 bytes on wire (216 bits), 27 bytes captured (216 bits)
⊟ IEEE 802.15.4 Command, Dst: Atmel_bf:ea:be:72:c5, Src: IeeeRegi_01:
  ⊞ Frame Control Field: Command (0xcc63)
    Sequence Number: 68
    Destination PAN: 0x1aaa
    Destination: Atmel_bf:ea:be:72:c5 (00:04:25:bf:ea:be:72:c5)
    Extended Source: IeeeRegi_01:30:0c:48:0d (00:50:c2:01:30:0c:48:0d)
    Command Identifier: Association Response (0x02)
  ⊟ Association Response
      Short Address: 0xffff
      Association Status: 0x01 (PAN Full)
    FCS: 0x5dbf (Correct)
```

Figure 2.1: Association request failure due to zbassocflood attack.

The consequence of a successful denial-of-service attack by *zbassocflood* is shown in Fig. 2.1. All available LR-WPAN network addresses have been allocated to devices that do not exist, as reported in the ‘PAN full’ line highlighted. The *zbassocflood* tool made repeated association requests using spoofed MAC addresses, exhausting the network address pool. No new legitimate devices are thus able to join the network. Recent works improve the quality and capabilities of LR-WPAN attack tools [GBM+12][RMW12], motivating the need for novel defenses.

Recent works investigate transceiver fingerprinting techniques to accurately verify wireless transceivers using unique physical features. However, utilization of unique physical features as fingerprints necessitates training sessions in which wireless transmissions from trusted devices are collected and differences useful for verification are identified. Alternatively, the physical layer (PHY) manipulation techniques presented herein verify actual transceiver *type* without the need for a training session or costly signal analysis equipment. All security techniques have limitations; the primary PHY manipulation framework limitation is that *inter-type* transceiver differentiation is possible while *intra-type* differentiation is not. Attack tool firmware such as KillerBee typically support a subset of all available transceiver types, so accurate *inter-type* verification is a highly valuable component for wireless defense-in-depth.

This chapter addresses four primary research goals:

1. Determine whether the PHY preamble manipulation framework is invariant with respect to received signal strength.
2. Identify PHY manipulations to discriminate among the eight transceiver types under test.
3. Demonstrate intrusion detection and transceiver-type fingerprinting techniques on real-world IEEE 802.15.4 networks.
4. Investigate PHY preamble manipulation framework potential for IEEE 802.11 networks.

These goals are addressed by transmitting packets with non-standard wireless preambles and identifying the resultant differences in packet reception among transceiver types.

This chapter is organized as follows. Section 2.2 describes recent work toward LR-WPAN security, including intrusion detection and RF fingerprinting. Section 2.3 explains the process by which wireless PHY headers are manipulated. Section 2.4 reports which preamble manipulations are useful for manufacturer discrimination. Section 2.5 describes and then demonstrates preamble manipulation for multi-factor device authentication incorporating physical layer (PHY) attributes. Section 2.6 demonstrates real-time intrusion detection using PHY manipulation. Section 2.7 presents a methodology for accurately identifying the true tranceiver type of a LR-WPAN device. Section 2.8 discusses possible attacks against PHY manipulation. Section 2.9 reports preliminary success with applying PHY preamble manipulation techniques to IEEE 802.11 devices. Section 2.10 concludes and proposes directions for future work.

2.2 Related Work

LR-WPANs fulfill critical functions in health care, automation, and smart energy systems. However, LR-WPAN devices are challenging to secure due to tight design constraints on cost, computing resources, and energy use. This section reviews recent work toward wireless intrusion detection, RF fingerprinting, and PHY manipulation.

2.2.1 Intrusion detection.

Conti et al. [CPMM11] designed a distributed protocol for detecting node replication attacks in wireless sensor networks. The proposed protocol was an improvement over earlier techniques in terms of energy efficiency and resilience to attack. Li et al. [LGZC13] addressed flood network attacks by malicious insiders though a distributed protocol of enforced rate limits. Some LR-WPAN nodes cannot support even relatively efficient bit-layer security protocols, so Yang et al. [YCTC13] detected and localized spoofing attackers by analyzing received signal strength (RSS).

2.2.2 RF fingerprinting.

Identification of device spoofing is the primary goal of RF fingerprinting. Key limitations of traditional RF fingerprinting techniques are that they require training and management of RF fingerprint models and sophisticated signal measurement hardware that is orders of magnitude more costly than the transceivers overseen. The resultant RF fingerprint database is also specific to a particular collection of devices in an environment. PHY preamble manipulation investigated herein exploits transceiver design characteristics that are invariant with respect to the environment. Instead of precise phase or frequency metrics, the only measurement required in this PHY manipulation techniques is whether or not the device under test transmits a reply to a corresponding request. Furthermore, PHY manipulation requires no signal measurement hardware and can even be performed with standard LR-WPAN transceivers.

Early research toward RF fingerprinting of LR-WPAN devices exploited spectral components of the transmission turn-on transient region for device classification. Rasmussen and Capkun [RC07] successfully classified CC1000 transceivers, while Danev and Capkun [DC09] classified CC1000 and CC2420 devices. Later, Danev et al. [DLCED10] demonstrated that signal replay attacks were effective against RF fingerprints of only five spectral components when the attacker was in the expected location of the spoofed device. More recent work by Ramsey et al. [RTM12] and Dubendorfer et al. [DRT12] demonstrated CC2420 device classification and verification using RF fingerprints of 99 to 297 time domain components. Successful replay attacks against lengthy RF fingerprints have not been demonstrated as of this writing.

One of the most frequently cited works on RF fingerprinting is the PARDIS system [BBGO08]. PARDIS utilizes a vector signal analyzer to classify the true identify of IEEE 802.11 transceivers from carefully analyzed wireless transmission features. As in [RC07][DC09][DLCED10][RTM12], transceivers from the same manufacturer and of the same type were correctly classified with greater than 99% accuracy. Actual implementation of a real-time wireless security system based on RF fingerprinting, however, remains a significant challenge. By limiting the wireless intrusion detection scope to the identification of transceiver hardware *type (e.g., manufacturer)* rather than *serial number*, it is demonstrated herein that PHY preamble manipulation can likewise achieve greater than 99% accuracy. Furthermore, this classification accuracy is achieved without the need for oscilloscopes or signal vector analyzers.

2.2.3 PHY manipulation.

Packet-in-packet frame injection attacks demonstrated by Goodspeed et al. [GBM$^+$11] rely upon wireless interference to corrupt a start of frame delimiter (SFD) of a first packet such that a second packet placed within the first is the one actually received. Instead of manipulating packet payloads and relying upon interference to corrupt PHY headers, this investigation manipulates and corrupts the PHY directly. Closely related to the work herein is that of Muntwyler et al. [MLLP12] wherein LR-WPAN communication was obfuscated through manipulated PHY spreading codes using a software-defined radio. Only devices with the ability to interpret the random spreading codes can receive such modified packets.

2.3 Wireless PHY Manipulation

2.3.1 Methodology.

LR-WPAN devices implement PHY features of the IEEE 802.15.4 standard in their transceiver hardware. This experiment utilizes a National Instruments (NI) Universal Software Radio Peripheral (USRP) NI USRP-2921 to explore PHY deviations from the standard. The NI USRP-2921 receives and stores standard (unmodified PHY) LR-WPAN packets as vectors of instantaneous In-phase and Quadrature (I/Q) measurements. The I/Q data array takes the interleaved form

$$[I_0\ Q_0\ I_1\ Q_1\ I_2\ Q_2\ ...\ I_n\ Q_n],$$

where n is the number of acquired samples. After manipulating the packet preambles in MATLAB, altered packets are replayed on the USRP. Beacon requests feature PHY payloads of 10 bytes. Given the standard IEEE 802.15.4 PHY header length of six byes (4-byte preamble + 1-byte SFD + 1-byte frame length), total transmission length is 16 bytes for beacon requests. Two O-QPSK symbols represent each byte for total transmission lengths of 32 symbols. Symbol duration is 16 μs [Soc06] resulting in total transmissions of 512 μs. The USRP collection rate is two million in-phase and quadrature (I/Q) sample pairs per second, sufficient for successful RF replays. The USRP streams RF recordings of standard packets to a Dell Precision M4500 laptop via a gigabit Ethernet cable.

Short packets for which receiver reply is compulsory are used to assess differences in transceiver implementations among manufacturers. LR-WPAN transceivers automatically respond to acknowledgment requests within 5 ms [Atm09]. Packets for which reply is compulsory include beacon requests and data requests with the acknowledgment flag set. Both beacon requests and data requests can be addressed to solicit a reply from one networked device at a time. If the receiver under test responds to packets with modified preambles, it means the receiver is able to correctly receive and interpret the messages. If the device never generates a reply in response to repeated messages, the packets must have been either corrupted by wireless interference or are completely uninterpretable by the receiver. Influence of wireless noise is mitigated by operating on IEEE 802.15.4 channel 26 (2.480 GHz), outside the spectrum of nearby IEEE 802.11g access points. Trial repetitions of 500 packets for each scenario further mitigate random wireless interference as a confounding influence.

Table 2.1 lists the eight transceiver types under test. Rather than refer to the eight transceiver types by their full device name throughout, the two letter abbreviations are used throughout Section 2.4. It is important to note that the internal radio components of XB, EM, and ST are all produced by Ember.

The standard IEEE 802.15.4 preamble consists of eight symbols, each representing the hexadecimal value `0x0`. In this chapter the standard preamble is manipulated in three ways: 1) the number of preamble symbols is decreased to fewer than eight, 2) the preamble symbol composition is altered, and/or 3) the Frame Length field is altered. All manipulations involve the removal or replacement of entire O-QPSK symbols. Individual symbols are removed from the preamble by replacing them with background noise of equal duration (16 μs) from elsewhere in the signal collection. Symbol replacement consists of copying symbols representing other binary values from elsewhere in the packet to the PHY header region. Validation of these processes is discussed in Section 2.3.2.

It is important to note that the author leveraged this I/Q manipulation technique solely due to familiarity with the experimental hardware. Any hardware with sufficient

Table 2.1: Eight transceiver types under test.

Manufacturer	Type	Abbrev.
Atmel	AT86RF230	AT
Digi International	XBP24CZ7PIS	XB
Freescale	MC13213	FS
Jennic	JN5148	JN
Microchip Technology	MRF24J40MA	MC
Silicon Labs	EM357	EM
STMicroelectronics	STM32W	ST
Texas Instruments	CC2420	TI

transceiver flexibility can be used to transmit packets with manipulated PHY preambles; this includes at least one standard LR-WPAN transceiver, the CC2420. Since the CC2420 can transmit any byte as the SFD in outgoing transmissions [Ins14], it can also be used to generate manipulated PHY headers for crafted packets. CC2420 packet reception characteristics, however, are not as flexible as explained in Section 2.8.

2.3.2 Process Validation.

Contents of the IEEE 802.15.4 PHY header, including the preamble, are stripped by the transceiver and are not accessible to higher layers of the network stack. It is therefore necessary to validate that symbol-wise preamble manipulations actually result in the intended changes.

In order to demonstrate that the preamble manipulation methodology is successful, arbitrary symbols are copied from the PHY header to the Frame Check Sequence (FCS) at the end of the transmission. In one such demonstration the SFD symbols `0xA7` are copied to the FCS in a beacon request (Fig. 2.2). This requires understanding of the data content

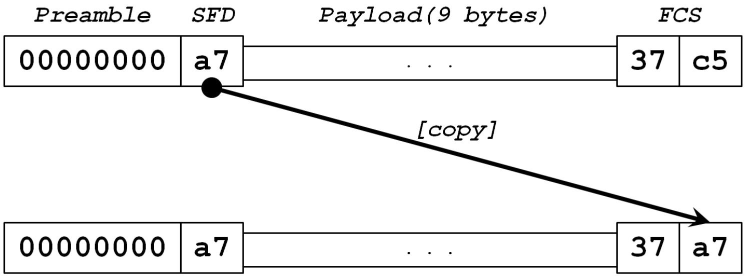
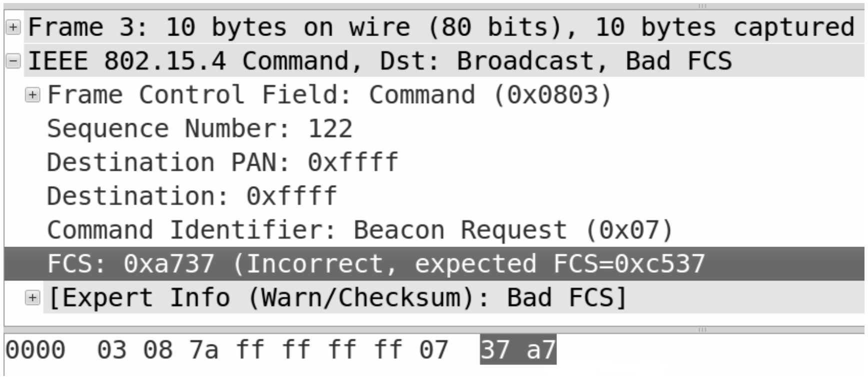


Figure 2.2: Conceptual illustration of MATLAB PHY manipulation featuring an altered FCS.



```
⊞ Frame 3: 10 bytes on wire (80 bits), 10 bytes captured
⊟ IEEE 802.15.4 Command, Dst: Broadcast, Bad FCS
  ⊞ Frame Control Field: Command (0x0803)
    Sequence Number: 122
    Destination PAN: 0xffff
    Destination: 0xffff
    Command Identifier: Beacon Request (0x07)
    FCS: 0xa737 (Incorrect, expected FCS=0xc537
  ⊞ [Expert Info (Warn/Checksum): Bad FCS]

0000  03 08 7a ff ff ff ff 07  37 a7
```

Figure 2.3: Wireshark capture showing successful PHY manipulation of the FCS.

of the beacon request and a geometric conception of where the SFD is within the transmission. Since the SFD immediately follows the preamble, knowledge of its exact location in the file to be manipulated and replayed by the USRP confirms that symbol-wise PHY manipulation of the preamble is also accurate. A beacon request is transmitted with this precisely-corrupted FCS to a packet sniffer for observation in Wireshark. The screenshot in Fig. 2.3 demonstrates that the FCS corrupts as expected. The correct FCS value of `0xc537` changes to `0xa737`, read in reverse byte order.

2.4 PHY Manipulation Results

2.4.1 Experiment Setup.

The IEEE 802.15.4 specification mandates a standard wireless preamble for every transmission. Manufactures design their receiver hardware to use this preamble for O-QPSK symbol synchronization of incoming transmissions [Soc06]. Exact receiver implementations of this process vary among manufacturers, however, due to the proprietary nature of hardware design. These subtle variations in hardware design are what this process seeks to leverage. Ultimately, the goal is to differentiate among transceiver manufacturers simply by observing how each device type responds to packets with PHY headers that deviate from the IEEE 802.15.4 standard.

The standard preamble consists of eight identical symbols, each representing the hexadecimal nibble `0x0`. Thousands of symbol-wise alterations to this standard are possible. The preamble can lengthen or shorten, and non-zero symbols can replace the standard zero symbols. This section reports packet reception rates for the eight transceiver types listed in Table 2.1 in response to PHY manipulation.

The eight transceiver types listed in Table 2.1 are powered one at a time, 0.5 meters from the USRP with an attached 3 dBi gain dipole antenna oriented vertically. An Atmel RZUSBstick [Atm09] reports mean received signal strength 0.5 meters from the transmitter of -61 dBm.

As each of the eight transceiver types are powered, the USRP transmits one beacon request per second for 500 consecutive seconds. The outcome of each beacon request transmission is a binomial process wherein one of two possible outcomes is realized; either the device under test correctly receives the packet and replies or it does not. Prior to each manipulated beacon request trial, the USRP transmits standard beacon requests without manipulation to establish that the transceiver under test is functioning normally. During subsequent trials outlined in Tables 2.2-2.8, the USRP transmits beacon requests with a particular manipulation (performed in MATLAB as described in Section 2.3) of the standard PHY.

2.4.2 Analysis of Device Replies.

The exact response percentages are not significant for the framework. The significant preamble manipulations are only those that never garner a response from a subset of transceiver types. Thus, if only one transceiver type ever responds to beacon requests with particular manipulated PHY, a single successful response (acknowledgment packet) is all that is necessary to correctly identify the true transceiver type of the device under test.

Tables 2.2-2.8 report transceiver response rates during the PHY manipulations under test. Reply rates for all transceiver types decline monotonically as additional trailing `0xF` nibbles replace the standard `0x0` nibbles within the preamble (Table 2.2). Dashes indicate that no replies are observed in response to the 500 manipulated beacon requests, and thus identify manipulations of greatest interest. For example, transceivers FS, JN, MC, and TI are never able to correctly receive beacon requests with `0xF` as the last nibble in incoming preambles (Table 2.2). Device reply rates and thresholds remain consistent even when received signal power at the transceivers under test decreases from -61 dBm (Table 2.2) to -76 dBm (Table 2.3). Received signal power is decreased by maintaining the transmitter/receiver distance and removing the transmission antenna from the USRP. These results demonstrate the salient nature of this PHY manipulation framework in response to varying signal-to-noise ratios. For all subsequent trials (Tables 2.4-2.8) received signal strength at the transceiver under test is -61 dBm.

Pilot studies established that packet replies in response to shortened preambles are equivalent to those in response to preambles with an equal number of leading `0xF` nibbles. For example, transceiver reply rates in response to a preamble of only four `0x0` nibbles were equivalent to reply rates in response to preambles with four leading `0xF` nibbles. For brevity, Table 2.4 is a packet reply report for both manipulation types. Here again, packet

Table 2.2: Packet reception versus modified preambles with trailing `0xF` nibbles (-61 dBm). Dashes represent zero packet reception.

Preamble	AT	XB	FS	JN	MC	EM	ST	TI
`0000000F`	100%	100%	-	-	-	100%	100%	-
`000000FF`	100%	79%	-	-	-	90%	100%	-
`00000FFF`	93%	65%	-	-	-	83%	17%	-

Table 2.3: Packet reception versus modified preambles with trailing `0xF` nibbles (-76 dBm).

Preamble	AT	XB	FS	JN	MC	EM	ST	TI
`0000000F`	100%	100%	-	-	-	100%	100%	-
`000000FF`	100%	86%	-	-	-	79%	100%	-
`00000FFF`	96%	78%	-	-	-	71%	20%	-

replies decrease monotonically as the preambles deviate more significantly from the IEEE 802.15.4 standard. Notably, none of the eight transceivers successfully receive packets with preambles that lack a `0x00` byte at the preamble tail. Only transceivers JN and TI successfully receive packets with the first three quarters of the standard preamble removed or replaced with `0xF` nibbles, so a successful reply in response to such packets narrows the true identify of the transceiver under test to JN or TI. The CC2420 transceiver ("TI") also serves as the radio core of more recent systems-on-a-chip, such as the CC2430 and CC2531. Analysis confirms that all such Texas Instruments chips follow equivalent response patterns.

Another interesting preamble manipulation is the replacement of one preamble byte with a false SFD byte `0xA7`. Table 2.5 reports transceiver reply rates in response to the

Table 2.4: Packet reception rates versus modified preambles with leading ones.

Preamble	AT	XB	FS	JN	MC	EM	ST	TI
F0000000	100%	100%	100%	100%	100%	100%	100%	100%
FF000000	92%	93%	96%	100%	100%	69%	100%	100%
FFF00000	-	-	68%	100%	63%	-	-	100%
FFFF0000	-	-	67%	100%	-	-	-	100%
FFFFF000	-	-	6%	100%	-	-	-	94%
FFFFFF00	-	-	-	73%	-	-	-	90%
FFFFFFF0	-	-	-	-	-	-	-	-

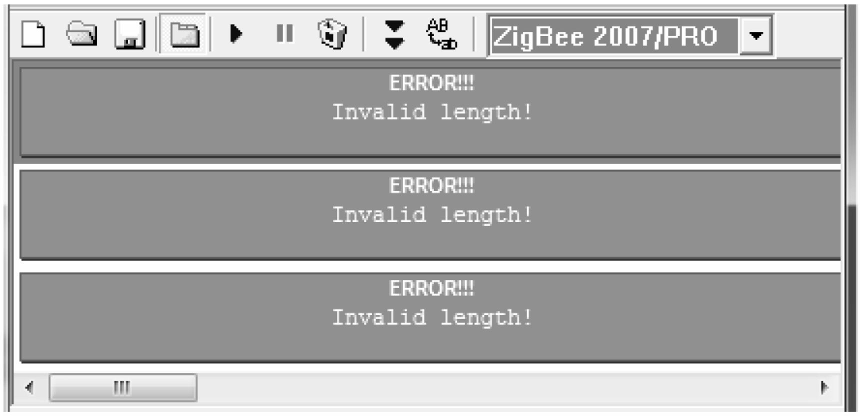
seven possible SFD insertions within a standard eight-nibble preamble. The false SFD presents significant challenges to packet reception, depending on its position within the preamble. The AT transceiver is the most resilient to this manipulation, replying to at least some of the beacon requests in every scenario. Contrastingly, packet reception ceases completely for transceivers XB, MC, and EM, and ST when the false SFD replaces the second byte or later.

Results confirm that the presence of a false SFD in the preamble causes packet reception failure due to misinterpreted length by analyzing packet reception with a Texas Instruments CC2531 packet sniffer. The screenshot in Fig. 2.4 is from the TI SmartRF Packet Sniffer interface. Three incorrectly received packets are in view, all modified with the 0x0000A700 preamble. For each observed packet the transceiver reports to the higher layers that the packet was of invalid length and is therefore indiscernible.

Table 2.6 reports packet reception rates in response to preambles with seven different 0xA nibble sequences. These preambles are non-standard, but do not cause disruption as significant as reported in Tables 2.4 and 2.5. These results provide valuable insight into

Table 2.5: Packet reception rates versus preambles with injected start frame delimiters.

Preamble	AT	XB	FS	JN	MC	EM	ST	TI
A7000000	100%	100%	100%	100%	100%	100%	100%	100%
0A700000	1%	53%	87%	83%	6%	59%	46%	100%
00A70000	5%	-	81%	20%	-	-	-	6%
000A7000	100%	29%	9%	-	-	7%	20%	80%
0000A700	100%	-	-	-	-	-	-	-
00000A70	100%	-	94%	-	-	-	-	-
000000A7	18%	-	-	-	-	-	-	-



ERROR!!!
Invalid length!

ERROR!!!
Invalid length!

ERROR!!!
Invalid length!

Figure 2.4: Invalid packet lengths reported by TI hardware in response to 0000A700 preambles.

the packet reception limitations of the eight transceivers. For example, from data in Tables 2.2-2.6 it becomes clear that the MC transceiver requires five trailing 0x0 nibbles in incoming packet preambles for successful message reception. Similarly, the TI and JN transceivers require two trailing 0x0 nibbles in incoming packet preambles for successful reception.

Table 2.6: Packet reception rates versus modified preambles with `0xA` nibbles.

Preamble	AT	XB	FS	JN	MC	EM	ST	TI
`AA000000`	100%	100%	100%	100%	100%	100%	100%	100%
`0AA00000`	1%	71%	88%	100%	89%	75%	74%	100%
`00AA0000`	7%	-	83%	100%	-	-	-	100%
`000AA000`	88%	61%	13%	100%	-	67%	48%	95%
`0000AA00`	91%	49%	88%	78%	-	39%	28%	90%
`00000AA0`	94%	100%	88%	-	-	100%	62%	-
`000000AA`	100%	72%	82%	-	-	73%	55%	-
`0000000A`	100%	84%	90%	-	-	75%	58%	-

In addition to the preamble and SFD, there is a third field in the IEEE 802.15.4 PHY header: Frame Length. The Frame Length field consists of the byte following the SFD, as briefly mentioned in Section 2.3. Maximum frame length as specified by IEEE 802.15.4 is 127 bytes, so the most significant bit in the Frame Length byte should be ignored. However, the eight transceivers under test are evenly split as to how this most significant bit is handled. Table 2.7 reports that when the Frame Length field is set to the standard 10_d, all eight transceivers receive and reply to incoming beacon requests. When the most significant bit of the Frame Length field is changed to a one, implying a frame length of 138_d, transceivers XB, EM, ST, and TI cease to reply. This Frame Length manipulation significantly compliments the preamble manipulations in Tables 2.2-2.6 and narrows the true hardware of the device under test to within one of four types.

Given the thousands of possible PHY manipulations, an exhaustive analysis of all permutations is beyond the scope of this work. However, in order to predict transceiver response rates to various manipulations, a decision tree model in MATLAB is generated using the data presented in Tables 2.2-2.7. PHY manipulations predicted to be useful by

Table 2.7: Packet reception rates versus frame length field in the PHY header (standard preambles).

Length	**AT**	**XB**	**FS**	**JN**	**MC**	**EM**	**ST**	**TI**
10_d	100%	100%	100%	100%	100%	100%	100%	100%
138_d	100%	-	100%	100%	100%	-	-	-

Table 2.8: Packet reception rates versus notable PHY header manipulations.

Preamble	**Length**	**AT**	**XB**	**FS**	**JN**	**MC**	**EM**	**ST**	**TI**
`0000A700`	138_d	100%	-	-	-	-	-	-	-
`0F070AFF`	10_d	-	16%	-	-	-	3%	2%	-
`0A07AA0A`	138_d	-	-	20%	-	-	-	-	-
`7A77A700`	138_d	-	-	-	100%	-	-	-	-
`FFFFFF00`	10_d	-	-	-	73%	-	-	-	90%

the model are then randomly selected, tested, and the results are incorporated to generate new and more accurate models. This process is repeated for 200 generations, at which point five PHY manipulations are identified as notable in their exclusivity of reception. These five PHY manipulations are shown in Table 2.8.

The first preamble manipulation in Table 2.8 is also reported in Table 2.5. The uniqueness of this preamble is made more robust by combining it with the Frame Length manipulation 138_d reported in Table 2.7. Only the AT transceiver is able to receive beacon requests using this manipulated PHY. Similarly, only the FS transceiver can receive beacon requests with the third PHY manipulation listed in Table 2.8. The JN transceiver, which cannot be differentiated from the TI transceiver through preamble manipulations alone, *can* be uniquely identified when an invalid Frame Length is used in conjunction with the fourth preamble listed in Table 2.8. Results also strongly suggest that the RF circuitry designs within the XB, EM, and ST transceivers are so closely related that the three chips are not distinguishable through PHY manipulation.

The following sections demonstrate the practical use of these results for PHY-augmented multi-factor authentication, wireless intrusion detection, and remote device type fingerprinting.

2.5 PHY-Augmented Authentication

2.5.1 Background.

Device authentication is a fundamental process in communication networks. Ostensibly immutable hardware addresses (e.g., MAC addresses) often serve as the "true" hardware identity. However, MAC address spoofing is straightforward with open source tools such as *macchanger* for IEEE 802.11 and *zbassocflood* for IEEE 802.15.4. Cryptographic credentials are another method for establishing device authentication, as long as the credentials have not been compromised. Multi-factor authentication relies upon two or more presentations by the device in question. For example, LR-WPANs may utilize both MAC address filtering and a private NWK key as two-factor authentication. The investigation and demonstration of PHY fingerprints as a third authentication factor is the subject of much recent work, as reviewed in Section 2.2.

A fundamental authentication exchange in IEEE 802.15.4 networks is the association request. Fig. 2.5 illustrates the association request message sequence between a device seeking to join the network and the LR-WPAN coordinator. The message sequence begins with an association request from the joining device. The association request includes the joining device's 64-bit (claimed) MAC address and is sent to the network coordinator's NWK address (`0x0000`). The coordinator's transceiver automatically replies with an acknowledgment while the coordinator begins determining whether or not the device may join the network. If the coordinator uses a MAC address white list for authentication, a network address will only be provided to a device presenting an approved MAC address. The coordinator also determines whether or not there are any unused network addresses in its address pool available for dissemination. While the coordinator performs these computations, the joining device waits a short period of time (e.g., one second) before requesting a response. After the response wait time the joining device sends a data request to the coordinator, generating another automatic acknowledgment. The final message is an association response from the coordinator. The association response either includes a valid network address assigned to the joining device, or it declines the association request (as in Fig. 2.1 for a full LR-WPAN).

2.5.2 Coordinator Authentication.

PHY manipulation is used to authenticate a coordinator based on its transceiver hardware type. As shown in Fig. 2.6, a USRP serves as the *joining device* that initiates the association request, a Freescale MC13213 serves as the network *coordinator*, Freescale MC13213 *end devices* form a functional network, and an Atmel AT86RF230 serves as a *traffic sniffer* to observe all message traffic.

In the first scenario PHY-augmented authentication is not active. The USRP transmits an association request to the coordinator using a standard PHY preamble, receives an acknowledgment, transmits a standard data request after a one-second wait, etc., completing the message sequence in Fig. 2.5. The wireless sniffer records the successful association request handshake as five packets displayed in Fig. 2.7. The coordinator informs the joining device that it has been accepted into the network and that is has been assigned the NWK address `0x796f`.

Next, a PHY-augmented authentication process is active. The joining device tests the true transceiver type of the coordinator to ensure it is of the expected type and not an

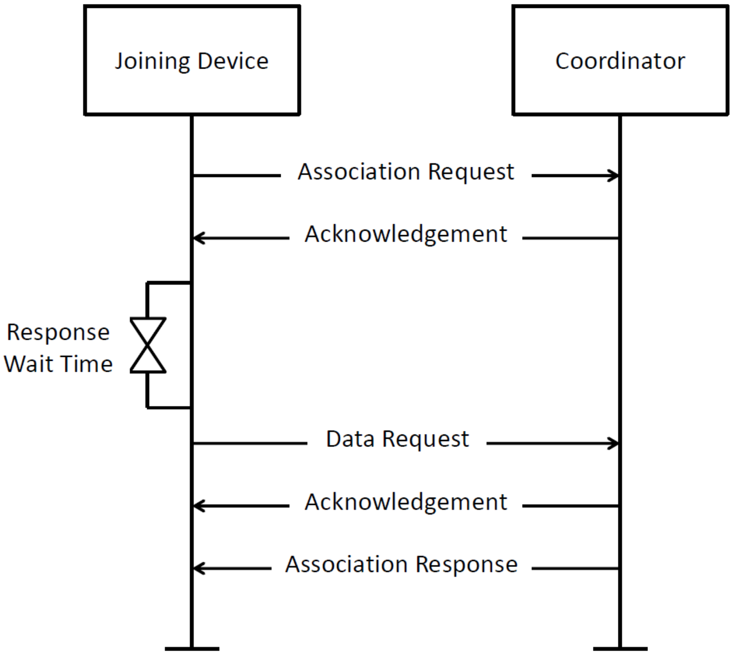


Figure 2.5: IEEE 802.15.4 association request message sequence.

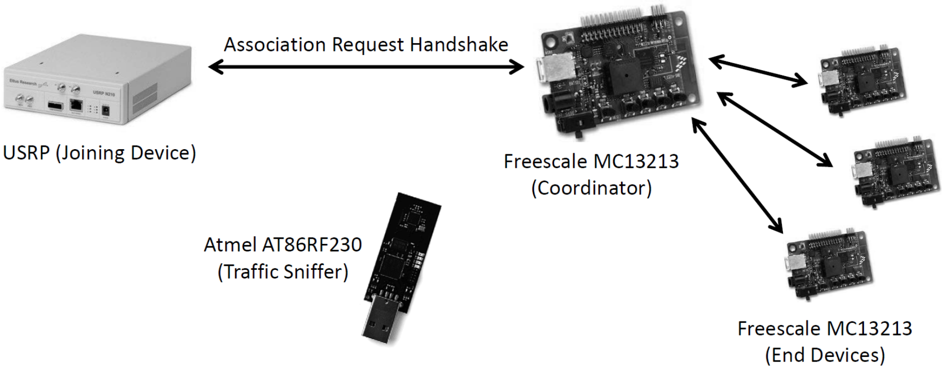
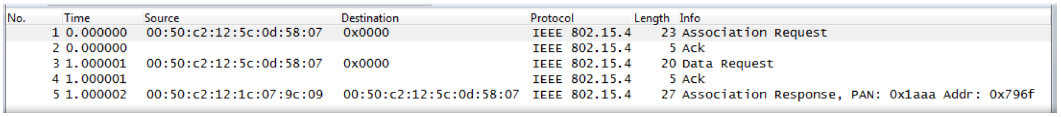


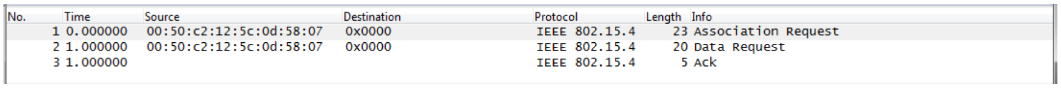
Figure 2.6: Demonstration setup (coordinator authentication).

impostor. In this scenario the trusted coordinator should have an AT86RF230 transceiver, but the device claiming to be the coordinator is actually a MC13213. The joining device authenticates the coordinator by modifying its transmissions such that only an



No.	Time	Source	Destination	Protocol	Length	Info
1	0.000000	00:50:c2:12:5c:0d:58:07	0x0000	IEEE 802.15.4	23	Association Request
2	0.000000			IEEE 802.15.4	5	Ack
3	1.000001	00:50:c2:12:5c:0d:58:07	0x0000	IEEE 802.15.4	20	Data Request
4	1.000001			IEEE 802.15.4	5	Ack
5	1.000002	00:50:c2:12:1c:07:9c:09	00:50:c2:12:5c:0d:58:07	IEEE 802.15.4	27	Association Response, PAN: 0x1aaa Addr: 0x796f

Figure 2.7: Successful association request sequence.



No.	Time	Source	Destination	Protocol	Length	Info
1	0.000000	00:50:c2:12:5c:0d:58:07	0x0000	IEEE 802.15.4	23	Association Request
2	1.000000	00:50:c2:12:5c:0d:58:07	0x0000	IEEE 802.15.4	20	Data Request
3	1.000000			IEEE 802.15.4	5	Ack

Figure 2.8: Failed association request sequence due to coordinator impostor.

AT86RF230 can receive them all. Specifically, its initial association request uses the `0000A700` preamble and its data request one second later has a `0000000A` preamble. Of the eight devices types under test, only the AT86RF230 can receive both of these packets. Fig. 2.8 shows the resulting message traffic. The coordinator cannot receive the modified association request, but it is observed by the AT86RF230 traffic sniffer. One second later the coordinator correctly receives the modified data request and its transceiver automatically replies with an acknowledgment, but no association response is sent because the coordinator is unaware of the initial association request. The joining device is then aware that the coordinator is not of the expected transceiver type.

This basic PHY-augmented authentication process is readily extendable to more complex handshakes including those using cryptography. If any part of the message handshake is not received by the intended recipient, the authentication fails. By crafting PHY manipulations for reception by the smallest possible number of transceiver types, PHY-augmented authentication of remote hardware is achieved.

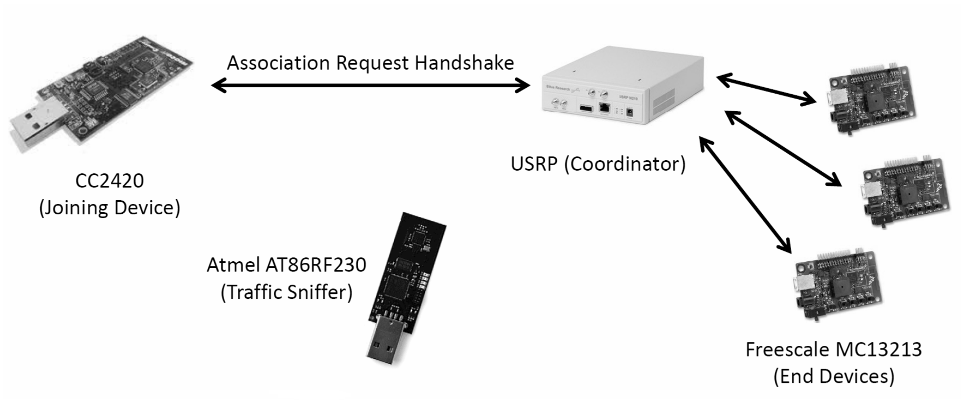


Figure 2.9: Demonstration setup (joining device authentication).

2.5.3 Joining Device Authentication.

PHY-manipulation-based authentication of joining devices mirrors the concept described in Section 2.5.2, but with the USRP serving as the coordinator and an impostor device attempting the join the LR-WPAN (Fig. 2.9). In order to implement this system in real-world hardware one additional feature must be implemented: sequence number management. When the USRP serves as the joining device, it determines the sequence numbers of its association request and data request packets, requiring the coordinator to reply with matching acknowledgment sequence numbers. When the coordinator is a standard device, sequence number management is handled automatically by its software. Since the Fig. 2.9 scenario requires PHY manipulation by the coordinator to authenticate the joining device, acknowledgment sequence numbers from the USRP must match those originating from the joining device. This is not possible though the signal replays performed previously. Instead, the USRP must be configured with actual LR-WPAN network functionality.

IEEE 802.15.4 network functionality is configured on the USRP using GNU Radio. IEEE 802.15.4 transmit and receive capabilities for USRP hardware via GNU radio have been steadily developed in recent years [Tha12] [SCTD09]. The implementation uses

USRP Hardware Driver (UHD) version 003.004.005 and GNU Radio version 3.6.1, running in Ubuntu 13.04 on a Dell Precision M4500 laptop. Preamble manipulations of outgoing transmissions from the USRP are accomplished through altering the *ieee802_15_4_pkt.py* script, wherein the preamble composition is specified.

The impostor CC2420 transmitter attempts to join the LR-WPAN by transmitting a standard association request. The USRP coordinator ensures that the joining device is *actually* a MC13213 transceiver by sending its acknowledgments with a `0A07AA0A` preamble and 138_d Frame Length. As reported in Table 2.8, this particular PHY manipulation is receivable by MC13213 transceivers, but not by any of the other transceiver types under test. As a result, the authentication handshake fails and the CC2420 is unable to join the network. Even if the impostor joining device "fakes" reception of the acknowledgment and proceeds with a data request anyway, the joining device will not be able to receive the PHY-manipulated association response which contains address and joining information. Furthermore, the impostor will not be able to correctly reply to any further PHY-manipulated interrogations or inquiries from the coordinator, and thus will not be advertised to the LR-WPAN as an authorized peer.

This authentication demonstration is a proof of concept that is straightforward to adapt to any IEEE 802.15.4-based network. Advanced adaptations of these techniques warrant future work, including obfuscated encryption key distribution in which only trusted end devices are able to receive sensitive data from the coordinator.

2.6 Wireless Intrusion Detection

Real-time intrusion detection is another promising use for preamble modification. In this experiment a sensor network utilizing Jennic JN5148 transceivers is placed throughout a building. Three intruder devices with alternate hardware are also placed throughout: a Digi International XBP24CZPIS, a Freescale MC13213, and a Silicon Labs EM357. Without PHY-based discrimination techniques, all transceiver types are indistinguishable

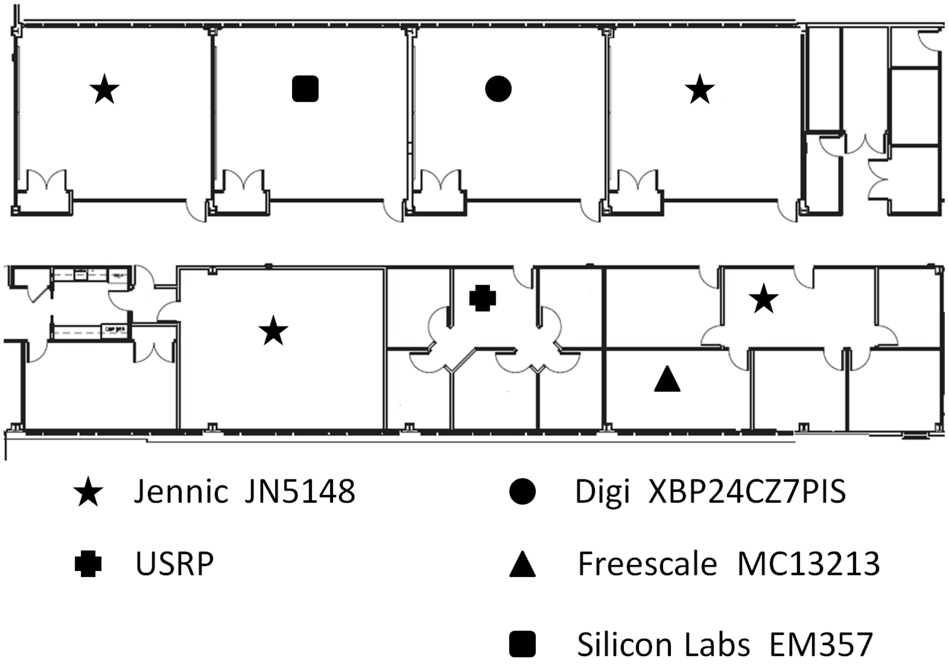


Figure 2.10: Device placement for intrusion detection demonstration within an office building.

to the network. A USRP serves as the intrusion detection system transmitter. Fig. 2.10 illustrates the demonstration topology. The four Jennic devices form a mesh sensor network reporting temperature, humidity, and light levels to a graphic display.

Jennic JN5148 transceivers cannot receive packets with the `000000AA` preamble modification. In this intrusion detection demonstration the USRP transmits a beacon request with a `000000AA` preamble modification once every few seconds at random intervals. The requests are dismissed as channel noise by the Jennic sensor network being protected and rarely interfere with legitimate traffic.

All three intruder devices receive packets with `000000AA` preambles. The IEEE 802.15.4 standard requires all full function devices to reply to beacon requests, so all intruding devices reveal themselves upon receiving each request.

The USRP transmits ten `000000AA` beacon requests within 70 seconds; ten replies are observed from the XBP24CZ7PIS and nine replies each are observed from the MC13213 and EM357. All three intruder devices reply to the first nine requests, each time revealing their differing hardware within milliseconds. No replies are generated from any of the JN5148 devices, as expected.

In an alternative protection scheme, devices already joined to the network are periodically audited with unicast packets. Any intruder device transceiver automatically replies to packets requesting acknowledgment, irrespective of software, thus immediately revealing its untrusted hardware. This alternative scheme is preferable for heterogeneous networks with multiple transceiver types because devices are audited individually.

These intrusion detection techniques compliment the authentication framework demonstrated in Section 2.5. If cryptographic credentials (such as a new NWK key) are distributed using the fourth PHY manipulation listed in Table 2.8, all three intruder devices will not have been able to intercept them. False keys could also be disseminated using `000000AA` preambles, and the three intruders would reveal themselves upon trying to use them.

2.7 Remote Device Type Fingerprinting

This section presents and demonstrates a methodology for classifying unknown transceiver hardware with high accuracy.

Tables 2.2-2.8 present preamble and Frame Length manipulations which can be used to accurately identify the transceiver type within an unknown or suspicious device. As a proof-of-concept, a classification decision tree is presented in Fig. 2.11. Note, this example is but one of the many possible decision trees possible and optimization of such trees warrants is own future work.

The process begins with successful reception of an ACK from the device under test using the standard IEEE 802.15.4 PHY. This establishes that the device under test is

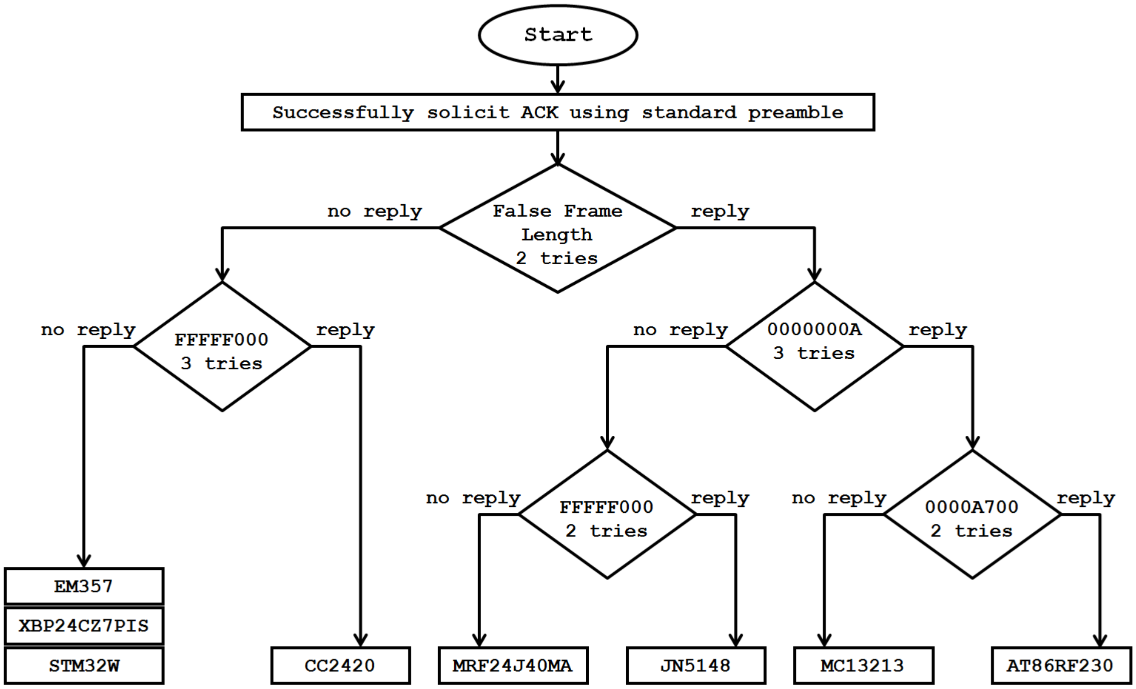


Figure 2.11: Example classification decision tree for IEEE 802.15.4 transceivers.

powered and responsive. Next, up to two packets requiring acknowledgments are transmitted to the device under test with the most significant bit of the Frame Length incorrectly manipulated to a one. Once an ACK is observed in response to a manipulated packet, additional packets are not necessary at that decision step in the classification tree. As shown in Table 2.7, a reply in response to this manipulation narrows the true transceiver to one of four types. Two tries are allowed during the first test (falsified Frame Length) to account for any interference or a dropped packet. Note that the number of tries can be arbitrarily increased at any decision point in the tree to suit desired accuracy and energy efficiency requirements. The number of tries suggested throughout Fig. 2.11 is designed to result in greater than 99% correct classification accuracy, assuming the experimentally-observed packet response rates reported in Section 2.4.

While this model is predicted to be effective mathematically, it must also be evaluated against real-world devices to demonstrate its effectiveness. Therefore, each of the eight device types are concealed one at a time in a cardboard box and the classification decision tree is used to determine the true transceiver type within. A NI USRP-2921 served as the PHY manipulation transmitter and responses are observed by an AT86RF230 packet sniffer 0.5 meters distant. Table 2.9 reports the classification results of the eight trials. The number of packets required includes the initial packet(s) used to confirm that the device under test is responsive. As expected, all transceivers under test are correctly identified.

2.8 Attacks against PHY Manipulation

This section discusses the potential for countermeasures to PHY manipulation intrusion detection and fingerprinting strategies. An attacker utilizing a trusted transceiver type against a PHY manipulation defense will not be detected. Of course, a physically-compromised trusted device will not be detected by sophisticated RF fingerprinting systems either. It is important to note that PHY manipulation is envisioned as a novel tool for wireless situational awareness and a powerful component within a layered defense, rather than a security panacea.

One tactic for defeating PHY manipulation is to determine whether or not the PHY of an incoming packet has been manipulated (and thus a test). If an attacker can determine that an incoming packet preamble has been manipulated, she can selectively ignore incoming packets to mirror the response behavior of a particular transceiver type. Such deception is reminiscent of firewall-based obstructions to *nmap* operating system fingerprinting [KS05]. Some arbitrary waveform generators and software-defined radios are able to provide this insight, which is not accessible to low-cost end devices. PHY information can be nontrivial to garner, even with a software-defined radio. For example, IEEE 802.15.4 GNU Radio scripts developed by Schmid et al. [SSS07] for the USRP do

Table 2.9: Classification results (eight real-world devices) using decision tree in Fig. 2.11.

Device	Packets Required	Accuracy
AT86RF230	4	correct
XBP24CZ7PIS or EM357 or STM32W	6, 6, and 6	correct × 3
MC13213	5	correct
JN5148	6	correct
MRF24J40MA	7	correct
STM32W	8	correct
CC2420	4	correct

not report incoming PHY headers. Instead, the entire incoming PHY header is abstracted away and a placeholder byte "0xff" is printed to the terminal when scripts based on the *uhd_cc2420_rxtext.py* are executed. Modification of the IEEE 802.15.4 GNU Radio scripts to reveal incoming preamble composition is presently beyond the technical abilities of the author. Nevertheless, there do not appear to be any technical reasons why such changes cannot be implemented in future work.

The alternative to analyzing incoming PHY headers is to physically alter the receiver characteristics of an attacking device to match those of the spoofed type. A review of the eight device type datasheets reveals few options for PHY customization through registry value changes. The most promising flexibility is offered by CC2420 hardware. By altering the CC2420 `SYNCWORD` register, the required synchronization sequence for incoming transmissions can change from `0x00A7` (as observed in Section 2.4) to `0x000A7`. That is, a CC2420 transceiver can made to require either two or three trailing `0x0` nibbles in incoming preambles. Fortunately, this minor flexibility is not sufficient to defeat the device classification tree in Fig. 2.11. In addition, no registry configurations alter the fact that CC2420 transceivers cannot receive packets with a manipulated Frame Length field. In summary, no evidence is found of significant threats to PHY manipulation posed by standard end devices.

2.9 Preliminary IEEE 802.11 Results

Transmission preambles are featured in numerous wireless standards, including proprietary sub-GHz protocols and IEEE 802.11 local area networks (LANs). The goal is to discover whether preamble modification techniques are also useful for classifying IEEE 802.11 device types. Preliminary results strongly suggest such classification is possible [Kul14].

IEEE 802.11 protocols are significantly more complex than IEEE 802.15.4. For simplicity the analysis begins by examining two transceivers operating at 2 Mbps on an IEEE 802.11b LAN: an Atheros AR928X and an Intel 4965AG. Standard long preambles consist of 128 bits modulated at 1 Mbps. Each preamble bit is thus 1 μs long. One bit is removed at a time from the preamble for up to ten bits, forming preambles ranging from 127 to 118 bits long. Although the IEEE 802.11b PHY, transmission frequency, and preamble durations all differ from IEEE 802.15.4, a similar methodology may be followed as that described in Sections 2.3-2.4 of signal recording, modification in MATLAB, and replay on the USRP. Wireshark monitors the wireless interface to observe incoming packets. Thirty modified packets are transmitted to the device under test for each preamble length and a 99% confidence interval is calculated for the mean. Fig. 2.12 reports packet reception probabilities for the Atheros AR928X, while Fig. 2.13 reports the equivalent probabilities for the Intel 4965AG.

Packet reception on the two devices vary significantly in response to shortened preambles. Reception rates decline abruptly on the Atheros transceiver, but decline somewhat linearly on the Intel transceiver. Notable preamble lengths are 124 bits and 120 bits. The Atheros transceiver does not receive packets with 124 bit preambles, while the

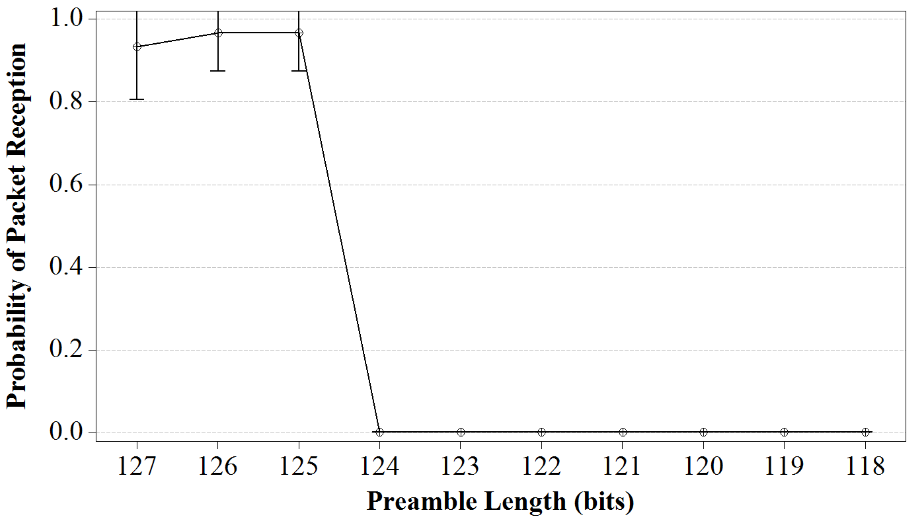


Figure 2.12: Packet reception versus preamble length on Atheros AR928X. Bars represent 99% CI based on thirty trials.

Intel transceiver receives approximately 50% of the packets. Neither transceiver receives packets with preambles shorter than 121 bits.

This investigation of IEEE 802.11 preamble manipulation is ongoing. The significantly longer preambles of IEEE 802.11 result in a wider range of possible bit-wise manipulations than for IEEE 802.15.4. Preliminary results demonstrate that preamble modification can augment bit-layer security processes for multiple wireless protocols.

2.10 Conclusion and Future Work

Radio frequency fingerprinting of wireless devices has attracted significant research attention in recent years. The ability to differentiate between wireless devices from subtle physical properties is a powerful defense against counterfeiting and network intrusion. Results herein demonstrate how preamble and Frame Length manipulation can be used for PHY-augmented device authentication, intrusion detection, and remote device type fingerprinting. Advantages to PHY manipulation include relatively low transmitter

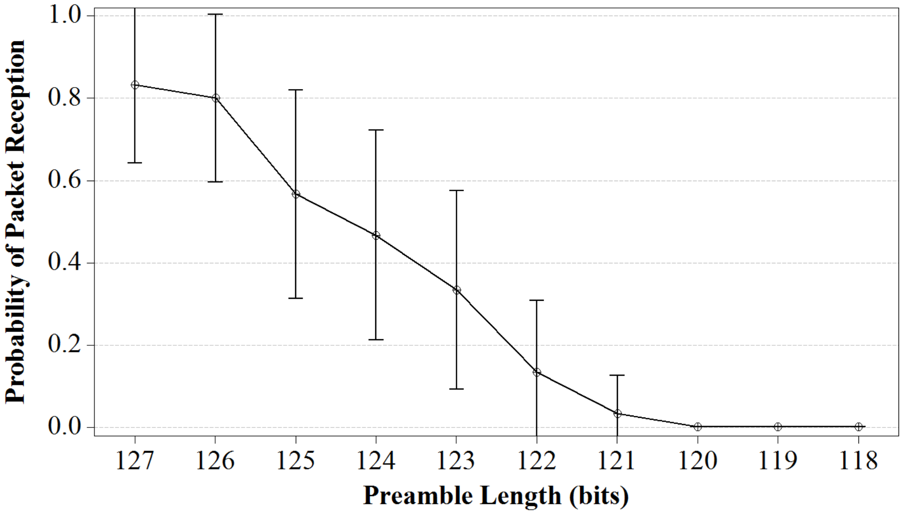


Figure 2.13: Packet reception versus preamble length on Intel 4965AG. Bars represent 99% CI based on thirty trials.

hardware cost, simplicity of implementation, and low computational complexity. The disadvantage of PHY manipulation over traditional RF fingerprinting is that it cannot be used to differentiate among devices within the same hardware class. However, bit-layer intrusion detection can compliment PHY manipulation to identify intruders using like-model (or even compromised legitimate devices) on the network.

Energy efficiency and throughput utilization are substantial considerations in many LR-WPANs. Preamble manipulation can be paired with traditional RF fingerprinting techniques to improve their efficiency. Intruders using untrusted hardware types can be rejected rapidly so that traditional RF fingerprints need only be generated and maintained for devices within the correct hardware class. Quantification of this benefit should be studied as traditional RF fingerprinting transitions from proof-of-concept to effective real-time systems.

In addition to preamble manipulation on IEEE 802.11 devices, other protocols that use preambles or similar synchronization features should be investigated. For example, midambles are a component of GSM networks that may also be exploited through PHY manipulation.

III. PHY Foundation for Multi-Factor LR-WPAN Node Authentication

3.1 Introduction

The primary limitation of the PHY manipulation framework outlined in Chapter 2 is that it cannot distinguish among transcievers from the same manufacturer and model number. In order to uniquely identify devices of the same type, RF fingerprinting must be used.

Previous work on RF fingerprinting for wireless sensor networks has exploited features within the signal turn-on transient region lasting approximately 125 ns [RC07][DC09][KK10], with percent correct classification of $\%C{\approx}70\%$ achieved using five relative amplitude features from ten CC1000 radios operating at 433 MHz [RC07]. Work in [DC09] shows improvement to $\%C{\approx}97\%$ using ten different CC1000 radios at distances of 15 cm. The use of three transient features is promising for classifying 2.4 GHz ZigBee node radios at distances of 40 meters [DC09]. RF fingerprinting based on differences in Automatic Gain Control (AGC) circuitry response have been less successful, with limited feature differences observed between six ZigBee devices from the same manufacturer at distances of 10 cm [KK10].

The IEEE 802.15.4 standard [Soc06] mandates use of a preamble based on 30 to 40 bits, with the actual length based on signal modulation type/order and operating frequency band. The 32-bit preamble for 2.4 GHz ZigBee nodes is 128 μs long, or approximately 1024 times longer than typical signal transients that have been exploited [RC07][DC09][KK10]. Preamble RF features have been effectively used to reliably differentiate IEEE 802.11a radios [STMM08][KTMR09] and IEEE 802.15.4 CC2420 radios [RTM12]. Relative to these earlier works, the work here 1) revalidates device differentiability of like-model 2.4 GHz CC2420 ZigBee transceivers at varying signal-to-noise ratio (*SNR*) using RF preamble fingerprint features, 2) extends validation to include differentiation of like-model sub-GHz (915 MHz) CC1000 radios, and 3) formalizes and validates a quantitative, statistic-based, pre-classification feature selection process that was conceptually introduced and qualitatively assessed in [RTM12]. Results herein are foundational to continued development of the envisioned PHY-MAC-NWK multi-factor authentication processor to augment current ZigBee bit-level authentication with information contained in RF fingerprint features. The passive monitoring approach is key to achieving backward compatibility with existing ZigBee devices.

3.2 Envisioned ZigBee Multi-Factor Framework

The concept of an "air monitor" that observes wireless network transmission characteristics to augment bit-layer security mechanisms is not new. However, many of the challenges associated with practical network integration have not been adequately addressed. The work here builds upon the air monitor concept and describes the envisioned integration into ZigBee LR-WPANs to improve security.

3.2.1 ZigBee Nodes & Topologies.

The IEEE 802.15.4 standard specifies two node classes, including: Full Function Devices (FFDs) and Reduced Function Devices (RFDs). FFDs are always actively listening on the network and are typically powered by a constant external power supply. RFDs are battery-powered and primarily operate in sleep mode, waking only to check for pending messages or periodic updates.

ZigBee uses FFD and RFD elements for three node classes: ZigBee Coordinator (ZC), ZigBee Router (ZR) and ZigBee End Device (ZED). The ZC and ZRs must be FFDs, while ZEDs can be either FFDs or RFDs. There can only be one ZC per LR-WPAN, and it is responsible for establishing the network, allocating NWK addresses, and routing traffic. The LR-WPAN fails without its ZC. ZRs extend the LR-WPAN physical range by routing messages between their child RFD ZEDs using multi-hop topologies, such as the Cluster Tree and Mesh topologies illustrated in Fig. 3.1 [RTM12].

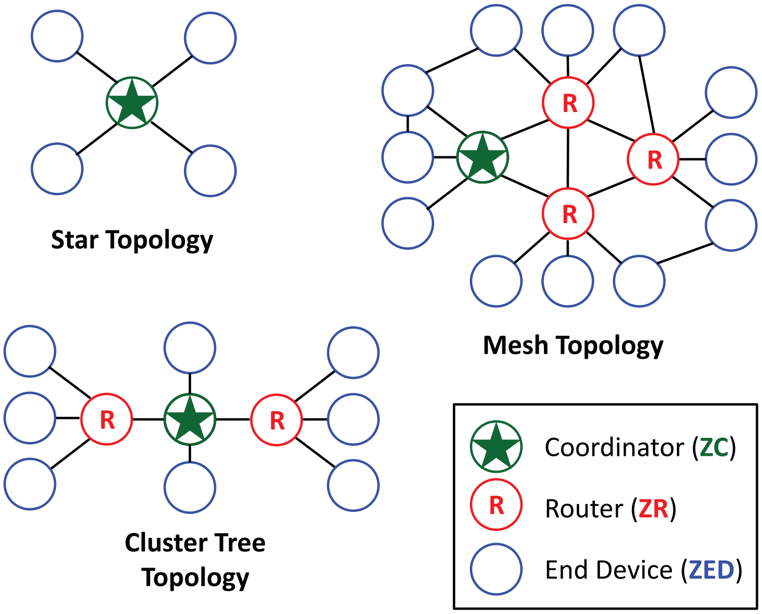


Figure 3.1: ZigBee LR-WPAN topologies [RTM12].

The Star topology is shown for completeness and does not support multi-hop communication. In a Cluster Tree topology, ZEDs have no children and can only communicate with the ZC and other ZEDs through their parent ZR. ZigBee Stack Profile 0x01 limits the number of children for each ZR to N_c=20, 6 of which can be ZRs. The ZigBee PRO specification (Stack Profile 0x02) increases this limit to N_c=254 children per ZR. Mesh topologies are only allowed using ZigBee PRO, and permit FFD ZEDs to communicate directly with one another.

3.2.2 Air Monitor Integration.

Air monitoring would be implemented using electronic devices that are separate from, but interfaced with FFD ZigBee devices to enable exchange of RF fingerprint information. A single air monitor would be sufficient for a Star topology if co-located with the ZC. For a Cluster Tree topology, an air monitor would be co-located with every ZR given that ZED communication is concentrated through its parent ZR. Mesh topologies pose significantly greater security challenges as memory overhead required for link key storage (confidentiality for every hop) can grow exponentially larger relative to cluster tree topologies. Air monitoring of large mesh topologies will be challenging for similar reasons.

Nodes are largely stationary in ZigBee applications used in Smart Energy, Building Automation, and Home Automation, which will simplify required air monitor coverage. However, ZigBee applications requiring mobile ZEDs, such as those commonly used in Health Care, pose significantly greater air monitoring challenges. Regardless, mobile ZED security must be addressed as they are inherently more vulnerable to physical attacks such as key extraction, theft, and tampering.

3.2.3 Air Monitor & Trust Center Integration.

ZigBee LR-WPANs under either security mode (standard or high) must appoint an FFD (usually the ZC) to serve as the Trust Center, recognized and trusted by all nodes on the LR-WPAN. The Trust Center is responsible for security and key management. A new node n* can only join the LR-WPAN if it receives permission from the Trust Center. Permission to join can be restricted by an access control list of valid MAC addresses. If n* presents a valid MAC address but does not know the network key, the Trust Center can transmit the key in plain text. ZigBee advocates assume that this window of vulnerability is “quite small and acceptable” [Ins12], but tools such as zbdsniff can endlessly sniff a LR-WPAN until such keys are intercepted [Wri09]. The proposed air monitor framework for ZigBee LR-WPANs would defend against active attacks such as fuzzing, associate request flooding, and packet injection by establishing a three-factor authentication process:

1. “Something you know” (NWK – Network keys);
2. “Something you have” (MAC – MAC address);
3. “Something you are” (PHY – RF fingerprint).

While the network keys and MAC addresses are vulnerable to current attacks, RF fingerprints from physical radio emissions are unique and technically infeasible to mimic.

In the Star topology in Fig. 3.1, the combined ZC/Trust Center receives feedback from its air monitor as to how well the current RF fingerprint from every incoming transmission matches the stored fingerprint profile established for the claimed sender. Thresholds for packet rejection must be tailored based on operational conditions to prevent undue denial of service. In Cluster Tree topologies, the ZRs only forward transmissions “cleared” as sufficiently well-matched by their respective air monitors. Air monitors maintain an evolving RF fingerprint profile of the devices assigned to its ZR to account for variations in environment and device operating characteristics. Sufficiently complex mesh networks require larger and more flexible RF fingerprint databases and air monitor placement.

An air monitor framework would be most valuable if every transmission is validated by the current RF fingerprint. This is because many exploits, such as replay attacks and packet injection, may be effective if a single malicious transmission is accepted as valid by the LR-WPAN. However, even fractional air monitor protection may mitigate active denial of service attacks such as associate request flooding.

Despite the challenges that must still be addressed before the envisioned air monitor framework is successfully implemented, the relatively low data rate ZigBee LR-WPANs, short, low-power transmission ranges, and an inherent limitation to N_c=20 or N_c=254 child devices per ZR makes them an ideal early candidate for emerging air monitor experimental research.

3.3 Background

3.3.1 Signal Collection Methodology.

An Agilent E3238S-based system [Tec09] serves as the RF Signal Intercept Collection System (RFSICS). All signal collections are down-converted to near-baseband, digitized using 12-bit analog-to-digital conversion and stored as complex in-phase and quadrature (I-Q) components for sub-sequent post-collection processing. Collection parameters included a sample frequency f_s=11.875 Msps and 4th-order Butterworth baseband filter bandwidth of W_{BB}=1 MHz. Signal collections included a total of N_P=1000 transmission preambles from N_D=7 CC2420 2.4 GHz IEEE 802.15.4 devices. Transceiver positioning was consistently maintained between collections in a Ramsey STE3000B RF test enclosure with RF-absorbent foam lining, 20 cm from a dipole antenna connected to the RFSICS input by a shielded cable.

Amplitude-based threshold detection was used with a leading edge value of T_D=−6.0 dB used to identify and extract individual burst transmissions from the multi-second RF collections. The approximate duration of experimentally collected preamble responses is 1536 samples (129 μs), which closely matches the 128 μs specification [Soc06]. The collection *SNR* for all bursts was SNR_C>50 dB.

3.3.2 Statistical Fingerprint Generation.

The statistical fingerprint (F) for a signal is derived from its instantaneous amplitude (a), phase (ϕ) and/or frequency (f) characteristics. More specifically, the sequences $a[n]$, $\phi[n]$, and/or $f[n]$ are generated from complex samples of the signal region of interest, centered (mean removal) and then normalized (division by maximum value) [STMM08][KTMR09]. Statistical fingerprint features are generated as variance (σ^2), skewness (γ), and/or kurtosis (κ) within specific signal regions. The regional fingerprint markers are generated by: 1) dividing each characteristic sequence into N_R contiguous, equal length sub-sequences, 2) calculating N_S=3 statistical metrics for each sub-sequence, plus the entire fingerprinted region as a whole (N_R + 1 total regions), and 3) arranging the metrics in a vector of the form

$$F_{R_i} = [\sigma^2_{R_i}\ \gamma_{R_i}\ \kappa_{R_i}]_{1\times 3}\ , \tag{3.1}$$

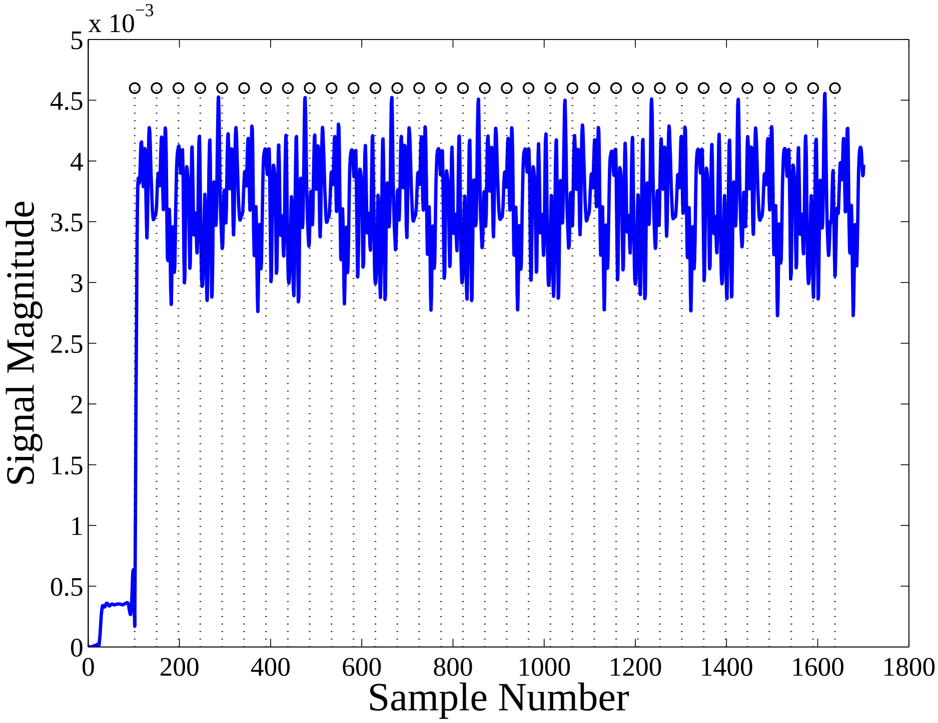


Figure 3.2: Representative burst preamble response with N_R=32 fingerprint sub-regions used for full-dimensional fingerprint generation [RTM12].

where $i = 1, 2, ..., N_R + 1$. The marker vectors from (3.1) are concatenated to form the composite characteristic vector for each characteristic and are given by

$$\mathbf{F}^C = \left[F_{R_1} \vdots F_{R_2} \vdots \ \ldots \ F_{R_{N_R+1}} \right]_{1 \times N_S(N_R+1)} . \tag{3.2}$$

If only one signal characteristic is used (a, ϕ, or f), the expression in (2) represents the final classification fingerprint. When all N_C=3 signal characteristics are used, the final RF fingerprint is generated by concatenating vectors from (2) according to

$$\mathbf{F}^C = \left[\mathbf{F}^a \vdots \mathbf{F}^\phi \vdots \ \ldots \ \mathbf{F}^f \right]_{1 \times N_S(N_R+1) \times N_C} . \tag{3.3}$$

While not optimally determined, empirical analysis revealed that $N_R = 32$ preamble sub-regions, or four regions per each of the eight repeated preamble sub-responses, was sufficient for establishing a proof-of-concept baseline. The subregions are illustrated in Fig. 3.2 for a representative preamble response.

3.3.3 MDA/ML Device Classification Methodology.

Statistical RF fingerprints are generated using (3.3) for device preamble transmissions from N_D=7 IEEE 802.15.4 CC2420 radios. The resultant RF fingerprints are classified here using a Multiple Discriminant Analysis, Maximum Likelihood (MDA/ML) process. MDA is a straightforward extension of the Fisher Linear Discriminant process when discrimination of more than two classes (devices) is required. MDA reduces the higher-dimensional input feature space with the goal of maximizing inter-class separation while reducing intra-class spread [DHS99]. For the N_C=3 class problems considered here, MDA/ML projects the multidimensional RF fingerprints into a 2-dimensional space. RF fingerprints are classified as being affiliated with one of N_C=3 possible classes based on Bayesian decision criteria using prior known probabilities, probability densities, and relevant costs associated with making a decision [HFT09]. For all results presented herein the associate costs are assumed equal for all classes.

The MDA/ML models are developed here using a K-fold cross-validation training process with K=5 to improve reliability. This value is consistent with literature which suggests values of K=5 and K=10 are sufficient [HFT09]. The best-performing model generated during the training process is subsequently used to generate testing results using a previously unseen collection of input features from each device. Unless otherwise noted, only classification testing accuracy results are reported in Section 3.4.

3.3.4 Pre-Classification Feature Selection.

This section formalizes a quantitative, statistic-based, pre-classification feature selection process that was conceptually introduced and qualitatively assessed in [RTM12]. In the aggregate, the assembled RF feature sets are effective for inter-device classification. However, RF fingerprinting work using other types of classifiers that provide a post-classification indication of feature relevance, e.g., Generalized Relevance Learning Vector Quantized Improved (GRLVQI) [RTO12] and Differential Evolution (DE)-Optimized Learning From Signals (LFS) classifiers [HWT11], have shown that individual RF fingerprint features do not generally contribute uniformly to overall classification performance. Although much less complex and more computationally efficient than the GRLVQI and LFS classifiers (for a given number of input feature dimensions), the MDA/ML classifier is inherently limited in that it provides no insight into feature relevance. The best characteristics of these various classifiers is desired, i.e, computational efficiency and the ability to use a minimum number of input feature components (dimensional reduction) to increase operational efficiency when fielded.

Intuitively, RF fingerprint components that exhibit maximal inter-device dissimilarity and minimal intra-device dissimilarity should be advantageous for MDA/ML classification. This form of pre-classification feature selection (input dimensionality reduction) is addressed here statistically by examining RF fingerprint components prior to MDA/ML classification. For a given signal type, the goal is to identify fingerprint components that possess statistical properties that are most advantageous for achieving reliable MDA/ML classification; the dimensional reduction goal is to reduce the RF fingerprint size (minimize N_F) while having minimal or tolerable impact on overall classification accuracy.

Empirical analysis has shown that RF fingerprint features extracted from collected signal preambles are non-normally distributed. Thus, nonparametric statistical analysis is appropriate and the Kolmogorov-Smirnov goodness-of-fit test (KS-test) is a suitable option for analyzing statistical feature differences. The KS-test is used here to quantify differences in cumulative distribution functions (CDF) between two sample distribution functions $S(X)$ and $T(X)$, e.g., RF fingerprint features from two devices. The numeric difference between $S(X)$ and $T(X)$ is defined as

$$S(X) = j/n\ , \tag{3.4}$$

where j is the number of points less than or equal to x and n is the total number of samples. If the sample $X_1, X_2, ..., X_n$ has been sorted in ascending order such that $X_1 \leq X_2 \leq ... \leq X_n$, the KS deviations K^+_{max} (maximum positive), K^-_{max} (maximum negative), and K_{max} (maximum absolute) can be computed by

$$K^+_{max} = n^{1/2} \max_{1 \leq j \leq n} \left\{ j/n - T(X_j) \right\}\ , \tag{3.5}$$

$$K^-_{max} = n^{1/2} \max_{1 \leq j \leq n} \left\{ F(X_j) - (j-1)/n \right\}\ , \tag{3.6}$$

$$K_{max} = max\{K^+_{max}, K^-_{max}\}\,. \tag{3.7}$$

The distribution functions of K^+_{max}, K^-_{max}, and K_{max} are known and tabulated, such that the null hypothesis is rejected when the computed statistics exceed critical values tabulated for the selected level of significance, i.e., alpha value. When KS-test results are presented as p-values, lower p-values indicate a more significant difference between data sets. When originally considered in the context of MDA/ML processing [RTM12], it was conjectured and qualitatively shown that lower KS-test p-value features possessed greater discrimination information and provided improved classification performance relative to higher p-value features.

The *qualitative* feature reduction assessment in [RTM12] provides a baseline for the *quantitative* process that is formalized next. Considering N_D=7 devices, all N_p=21 possible unique pairwise device combinations are considered, i.e., N_p=21 KS-tests (α=0.1) conducted at SNR=8.0 dB for N_F=(N_R + 1=33)×(N_S = 3)×(N_C = 3)=297 features. Fig. 3.3 shows summed p-values for the corresponding KS-tests. As visually indicated by the collection of lowest p-values, phase features collectively possess greater

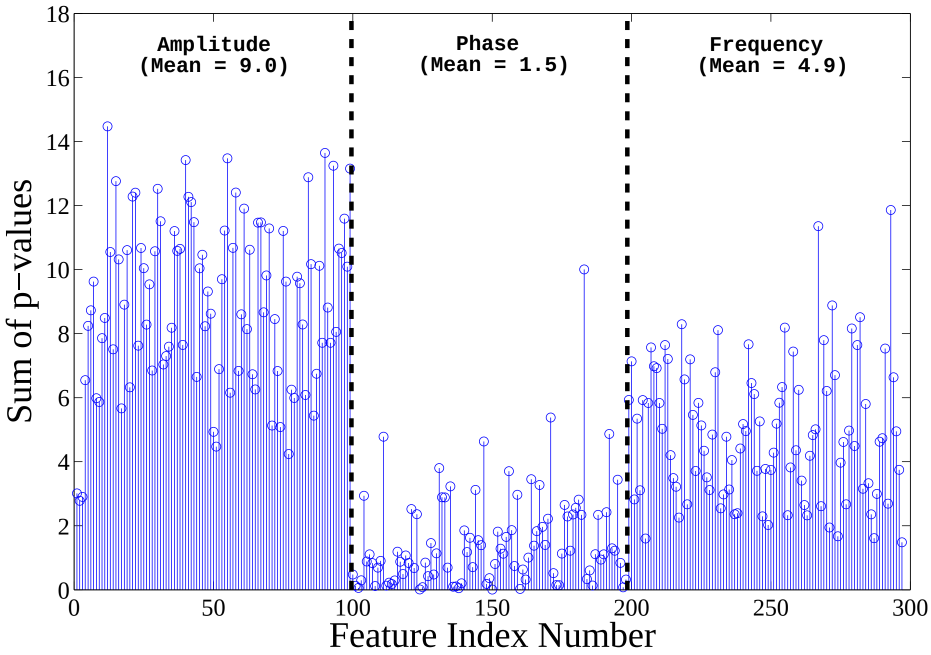


Figure 3.3: Sum of N_p=21 pairwise KS-test p-values for each fingerprint component of full-dimensional (N_F=297) features at SNR=8.0 dB [RTM12].

variation when compared to frequency features, which in turn possess greater variation when compared to amplitude features. The indicated robustness of phase features in Fig. 3.3 is consistent with what has been reported in earlier related work [RTM12].

As formalized here, a quantitative pre-classification KS-test feature selection process can be used to identify and select a most relevant, arbitrary-length l, subset of the full-dimensional RF feature set **F** prior to MDA/ML classification. The process is completed in six steps as follows:

1. Generate a full-dimensional (N_F) feature set using (3.1) through (3.3) for N_P preambles from each of the N_D devices to be classified. The preamble responses are combined with like-filtered Additive Gaussian White Noise (AWGN) to establish the desired SNR.

2. Conduct pairwise two-sample KS-tests using the N_F dimensional feature sets (n=N_P) between every two devices under test (α=0.1). Form a matrix of resultant p-values with dimension $N_D \times N_F$.

3. Rank-order the p-values in each row in non-increasing order.

4. Determine the feature index number corresponding to the lowest p-value in every row. If not already present in subset S, add the feature index number.

5. Check size[S] and remove indices corresponding to the largest p-values if size[S] > l.

6. Repeat Step 4 and Step 5 until S is of length l.

The quantitative pre-classification feature reduction process is applied for the N_D=7 devices under test. Results are presented in Fig. 3.4, Fig. 3.5, and Fig. 3.6 for the highest-ranked (lowest p-values), mid-ranked (middle p-values), and lowest-ranked (highest p-values) N_F=33 features in Fig. 3.3, respectively. Note that the vertical axis scales are different for visual clarity. As shown in Fig. 3.4, the most relevant (Top) N_F=33 components are exclusively derived from phase features, reinforcing the qualitative assessment made from Fig. 3.3 and results in [RTM12] that phase-derived features tend to be most effective for MDA/ML classification. Likewise, the lowest-ranked (Bottom) N_F=33 features in Fig. 3.6 are predominantly amplitude-derived which is again consistent with the qualitative assessment made from Fig. 3.3 and results in [RTM12].

3.4 2.4 GHz CC2420 Device Assessment

The IEEE 802.15.4 specification mandates the use of a Synchronization Header (SHR) containing a Preamble and Start-of-Frame Delimiter (SFD) sequence for all transmission bursts. Although the entire SHR can be used for generating RF fingerprints, empirical analysis revealed that inclusion of the SFD response did not significantly

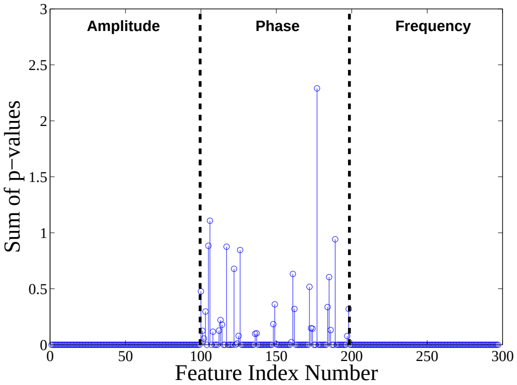


Figure 3.4: Highest-ranked (lowest *p*-values) N_F=33 features in Fig. 3.3.

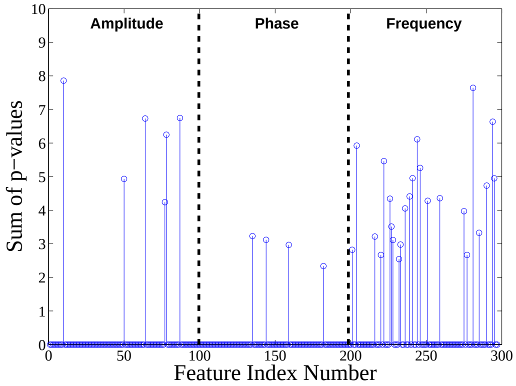


Figure 3.5: Mid-ranked (middle *p*-values) N_F=33 features in Fig. 3.3.

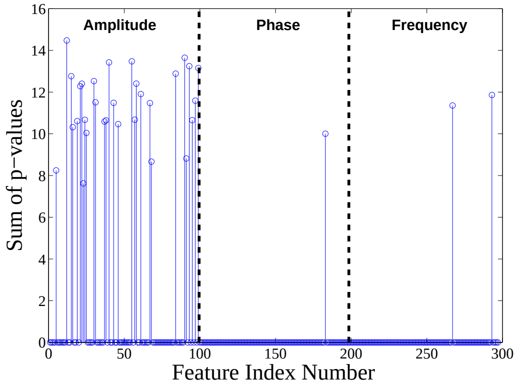


Figure 3.6: Lowest-ranked (highest *p*-values) N_F=33 features in Fig. 3.3.

improve MDA/ML classification accuracy. Additional analysis revealed that features based on power spectral density (PSD) significantly underperformed relative to features based on the instantaneous a, ϕ, and f time-domain responses considered here.

MDA/ML inter-device classification results were generated for all N_{Prm}=35 possible 3-class permutations of N_D=7 ZigBee devices. Classification experiments were conducted using N_P=1000 independent preamble responses (500 each for MDA/ML training and testing) and N_{Nz}=5 Monte Carlo noise realizations per preamble response at each *SNR*; a total of N_{Tst}=(500 Preambles)×($N_{Nz} = 5$)=2500 independent classification decisions per device in each 3-device trial. This large number of trials reduced the mean error bars to within the vertical extent of the plotted data markers. Therefore, trial mean error bars are intentionally omitted in all results plots to enhance visual clarity.

3.4.1 Full-dimensional RF Fingerprinting Performance.

Full-dimensional RF fingerprints include features based on N_C=3 signal characteristics (a, ϕ, and f), N_S=3 statistical fingerprint features (σ^2, γ, κ), and $N_R + 1$=33 regions, for a total fingerprint F comprised of N_F=297 RF fingerprint features as given by (3.3).

Fig. 3.7 shows aggregate full-dimensional classification (testing) performance accuracies for N_{Prm}=35 device permutations at *SNR*∈[0 20] dB. The cross-permutation average is shown as filled circle markers. Considering an arbitrary average percent correct classification of *%C*≥90% as a reasonable benchmark for assessing the potential contribution of PHY information to an overall PHY-MAC-NWK multi-factor authentication solution, the full-dimensional feature set successfully achieves the *%C*=90% benchmark at *SNR*=8.0 dB and would be suitable for a PHY-MAC-NWK integration.

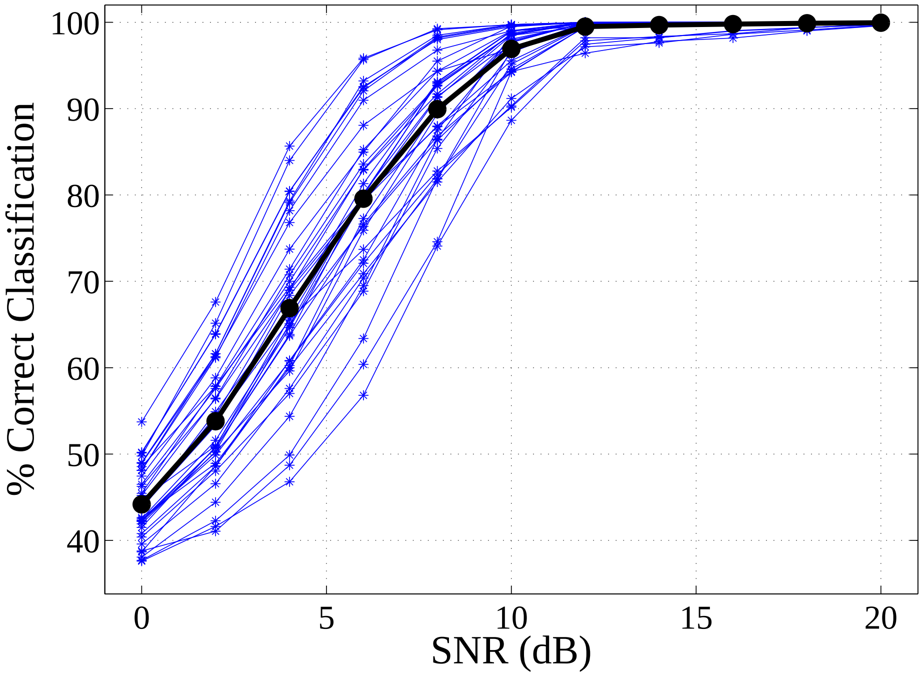


Figure 3.7: 2.4 GHz CC2420 MDA/ML classification: Full-dimensional (N_F=297) feature set and N_{Prm}=35 3-class permutations. Permutation average shown as filled circle markers [RTM12].

3.4.2 Reduced-Dimensional Qualitative Feature Selection.

While full-dimensional RF fingerprinting is effective, qualitative visual assessment of KS-test p-values in Fig. 3.3 reveals significant differences among RF fingerprint features derived from the instantaneous $a[n]$, $\phi[n]$, and $f[n]$ sequences. Classification results are presented here using dimensionally-reduced feature sets containing approximately 33% of the full-dimensional features (N_F=99 of 297). This is done by evaluating classification performance using only amplitude (Amp-Only), phase (Phz-Only) and frequency (Frq-Only) feature subsets.

Fig. 3.8 provides an overlay of the average cross-permutation classification performance for the full-dimensional feature set taken from Fig. 3.7 and the

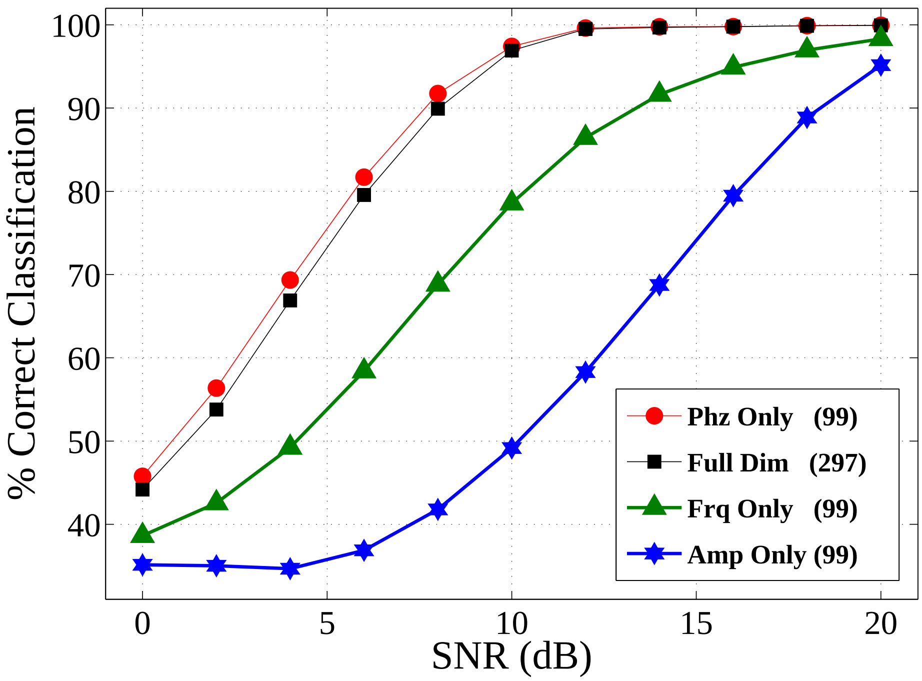


Figure 3.8: 2.4 GHz CC2420 MDA/ML classification: Permutation averages for full-dimensional (N_F=297) and reduced-dimensional (N_F=99) feature sets based on *qualitative* assessment of p-values in Fig. 3.3. [RTM12].

reduced-dimensional Amp-Only, Phz-Only) and Frq-Only feature sets for $SNR \in [0\ 20]$ dB. Again considering the arbitrary $\%C \geq 90\%$ benchmark for assessing PHY contribution to an PHY-MAC-NWK multi-factor authentication solution, the reduced-dimensional (N_F=99 of 297) Phz-Only feature sets provide near equivalent performance as the full-dimensional set and would perform reliably for $SNR \geq 8.0$ dB.

However, the reduced-dimensional Phz-Only feature set has an inherent advantage over the full-dimensional set in that it would only require calculation and processing of one-third the number of RF fingerprint features. This is a significant advantage when considering computational efficiency, speed and storage requirements for fielding an operational PHY-MAC-NWK authentication system.

With regard to the other reduced-dimensional results in Fig. 3.8, average cross-permutation classification accuracy using Amp-Only feature sets significantly under-performs all others as predicted by having the highest KS-test p-values in Fig. 3.3. The average cross-permutation classification accuracy for Frq-Only features falls between that of the Amp-Only and Phz-Only feature sets and achieves the arbitrary $\%C \geq 90\%$ benchmark for $SNR \geq 14.0$ dB.

3.4.3 Reduced-Dimensional Quantitative Feature Selection.

The statistical, quantitative p-value assessment process enables identification and selection of a most relevant subset of the full-dimensional features. The process is demonstrated here using reduced-dimensional feature sets comprised of a specific number of most relevant features. In this case, the specific number of $N_F=33$ features was chosen based on using a sufficient number of the highest ranked (lowest p-value) features required to achieve near identical classification accuracy as the full-dimensional set.

Fig. 3.9 provides an overlay of average cross-permutation classification accuracies using the top-ranked, middle-ranked, and bottom-ranked $N_F=33$ RF fingerprint features from Fig. 3.3 for the CC2420 devices. As indicated, the top-ranked $N_F=33$ feature set achieves the arbitrary $\%C \geq 90\%$ benchmark for $SNR \geq 9.0$ dB-actually less than a 1.0 dB difference relative to using full-dimensional and Phz-Only feature sets. The dimensional reduction versus performance trade-off is quite notable when considering that $N_F=33$ of 297 full-dimensional and $N_F=33$ of 99 Phz-Only features represent dimensional reductions of approximately 88% and 66%, respectively.

While the qualitative assessment in Section 3.4 was insightful, quantitative results here clearly indicate that the proposed KS-test approach to pre-classification feature selection, as formalized in section Section 3.4, is indeed valid and provides an effective means for identifying and rank-ordering the most relevant features for MDA/ML classification.

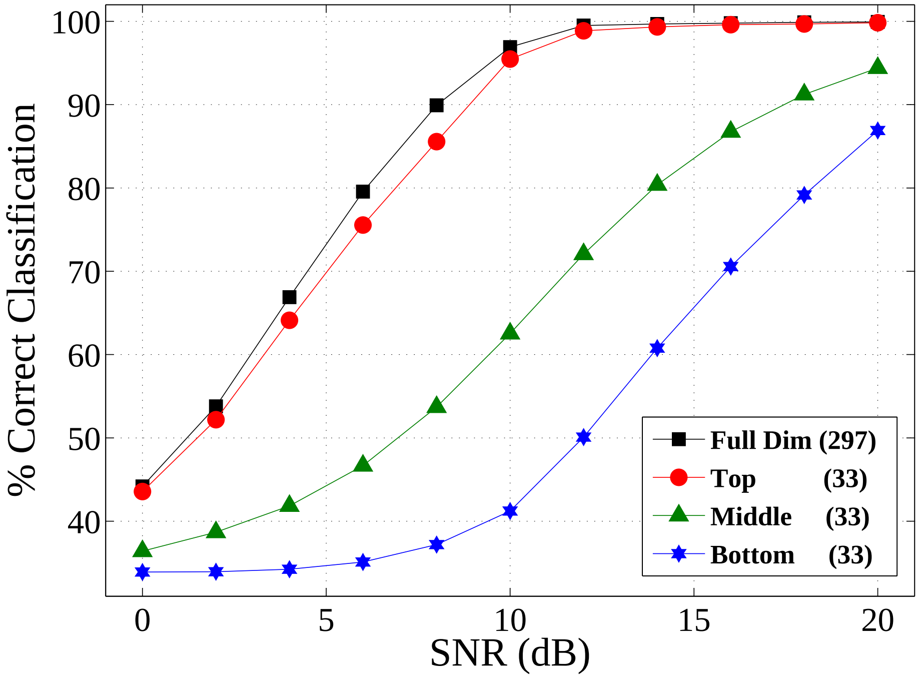


Figure 3.9: 2.4 GHz CC2420 MDA/ML classification: Permutation averages for full-dimensional (N_F=297) and *quantitatively* selected top-ranked, middle-ranked, and bottom-ranked reduced-dimensional (N_F=33) feature sets.

3.5 915 MHz CC1000 Device Assessment

The preamble-based fingerprinting and dimensional reduction techniques in Section 3.4 for CC2420 ZigBee devices was repeated to assess sub-GHz device authentication. This was done with CC1000 devices which use a proprietary PHY protocol and provide service in the Industrial, Scientific, and Medial (ISM) frequency bands: f_{ISM}=315, 433, 868, or 915 MHz. The CC1000 devices are used in home automation and Automatic Meter Reading (AMR) applications, each of which can be supported within the Smart Grid [Ins09].

As in Section 3.4, inter-device classification results and feature dimensional reduction assessment is demonstrated for N_D=7 CC1000 transceivers operating at

f_c=915 MHz. Given operational differences, the CC1000 collection parameters differed from those used for the 2.4 GHz ZigBee devices and included a final sample frequency of f_s=59.3755 Msps and a 4th-order Butterworth baseband filter bandwidth of W_{BB}=100 KHz. The number of collected preambles processed per device remained at N_P=1000.

The preamble length for the Frequency-Shift-Keyed (FSK) CC1000 PHY waveforms can be configured based on the intended transmission range. For the N_D=7 CC1000 devices used here, the experimentally observed preamble responses lasted approximately 4,000 μs. Initial exploratory analysis showed that RF features extracted from the entire preamble region did not contain sufficient information to perform reliable inter-device classification. However, features extracted from the first 253 μs of the preamble proved to have sufficient information and were adequate for initial proof-of-concept assessment. Fig. 3.10 shows the CC1000 preamble response and the N_R=15 sub-regions for RF feature extraction. The amplitude-based t_D=−6.0 dB burst detection point corresponds to sample number 501, while the preceding 500 samples are collected background noise. Relative to the CC2420 ZigBee devices, the CC1000 devices use a simpler PHY waveform structure and operate at lower data rates (less than 20 Kbps). Subsequent results will show that this increases the inter-device differentiation challenge and that future work remains to optimize parameter selection for associated signal collection, burst detection, and RF fingerprint generation.

3.5.1 Full-Dimensional RF Fingerprinting Performance.

Full-dimensional RF fingerprints include features based on N_C=3 signal characteristics (a, ϕ, and f), N_s=3 statistical fingerprint features (σ^2, γ, and κ), and $N_R + 1$=16 regions, for a total fingerprint F comprised of N_F=144 RF fingerprint components as given by (3.3). Fig. 3.11 shows the aggregate full-dimensional

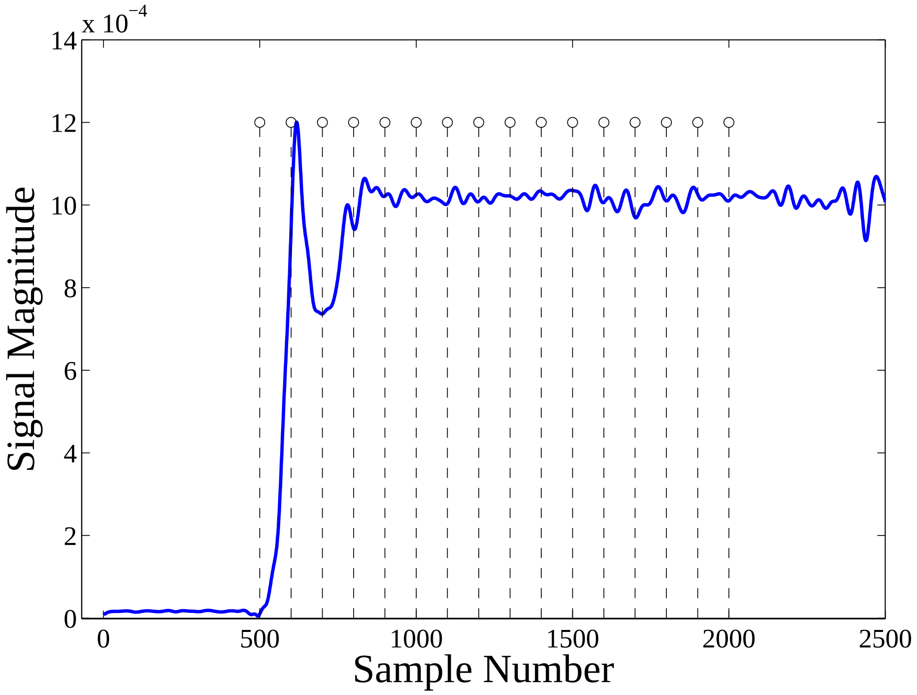


Figure 3.10: Provisioning of CC1000 burst response into N_R=15 sub-regions.

classification accuracies for N_{Prm}=35 device permutations at $SNR \in [0\ 20]$ dB. The cross-permutation average is shown as filled circle markers.

As shown in Fig. 3.12, the mean classification accuracy achieves the arbitrary $\%C \geq 90\%$ benchmark for $SNR \geq 18.0$ dB. While classification performance is lower than achieved with 2.4 GHz ZigBee devices in Fig. 3.7, these results do illustrate the potential for CC1000 inter-device differentiation using RF fingerprint features. Furthermore, the $\%C \geq 90\%$ benchmark was arbitrarily introduced in Section 3.4 for convenience and to enable comparative performance assessment. As planned development of the envisioned PHY-MAC-NWK authentication framework continues, it may be shown that $\%C \geq 80\%$ performance is a sufficient PHY contribution to the overall multi-factor solution. If so, the

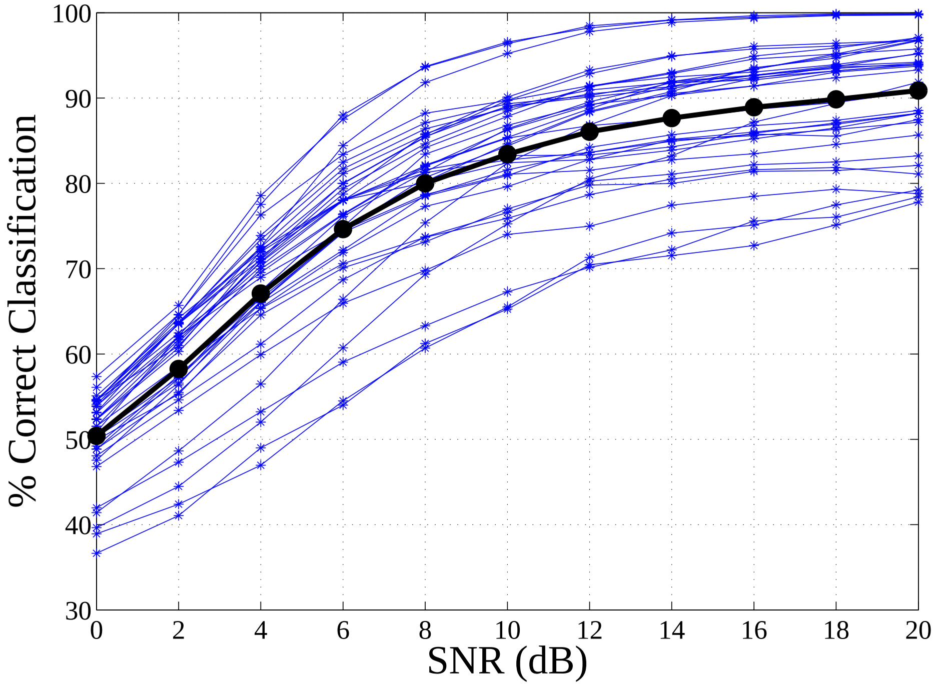


Figure 3.11: 915 MHz CC1000 MDA/ML classification: Full-dimensional (N_F=144) performance for N_{Prm}=35 3-class permutations. Permutation average shown as filled circle markers.

current technique is sufficient given that %C=80% CC1100 discrimination is achieved for SNR≥8.0 dB.

3.5.2 Reduced-Dimensional Quantitative Feature Selection.

Final CC1000 results include reassessment of the quantitative pre-classification feature selection process in Section 3 using reduced-dimensional sets having 50% of the full-dimensional features. The top-ranked and bottom-ranked N_F=72 of 144 features were used to produce results in Fig. 3.12 and Fig. 3.13, respectively. Classification performance for N_{Prm}=35 3-class permutations is provided along with average classification. Two notable conclusions can be drawn, including: 1) average top-ranked feature results in Fig. 3.12 are statistically equivalent to full-dimensional results in Fig. 3.11 for all *SNR*

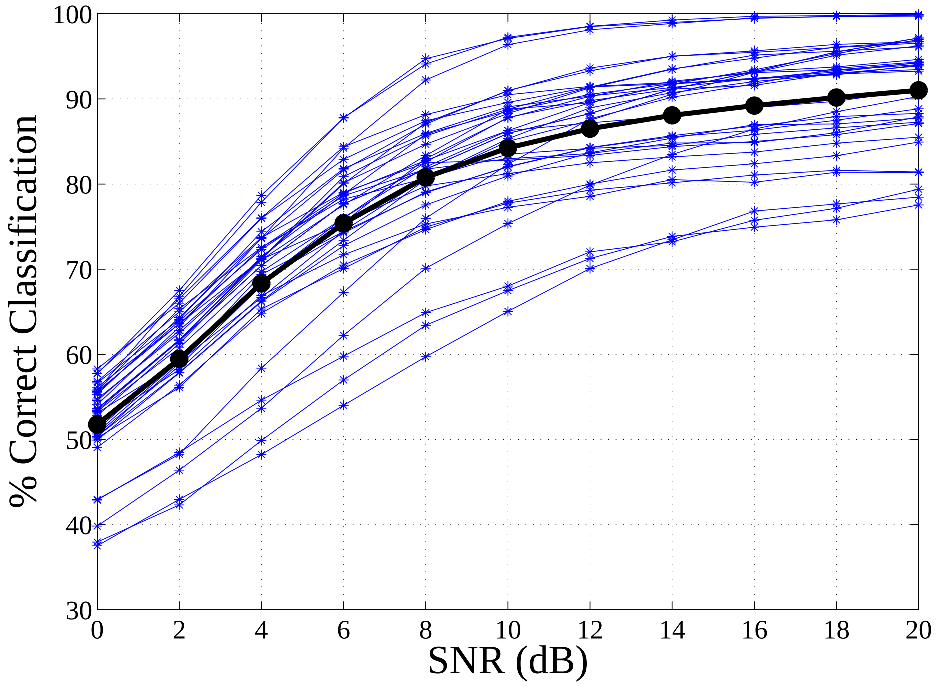


Figure 3.12: 915 MHz CC1000 MDA/ML classification: Quantitatively selected *top-ranked* reduced-dimensional (N_F=72 of 144 features) performance for N_{Prm}=35 3-class permutations. Permutation average shown as filled circle markers.

considered, and 2) average bottom-ranked results in Fig. 3.13 are statistically poorer by an average of 20% or more for $SNR \geq 8.0$ dB.

3.6 Quantitative Feature Selection Assessment

Reduced-dimensional results in Fig. 3.9 for CC2420 devices (N_F=33 of 297 features), and Fig. 3.12 for the CC1000 devices (N_F=72 of 144 features), clearly demonstrate the effectiveness of the KS-test *p*-value feature selection method developed in Section 3.3. While successful, it is important to note that the number of reduced-dimensional features was selected non-optimally with a goal of achieving near-identical performance using both reduced and full-dimensional sets. The relationship between selected reduced-dimensional features and the full-dimensional sets is illustrated

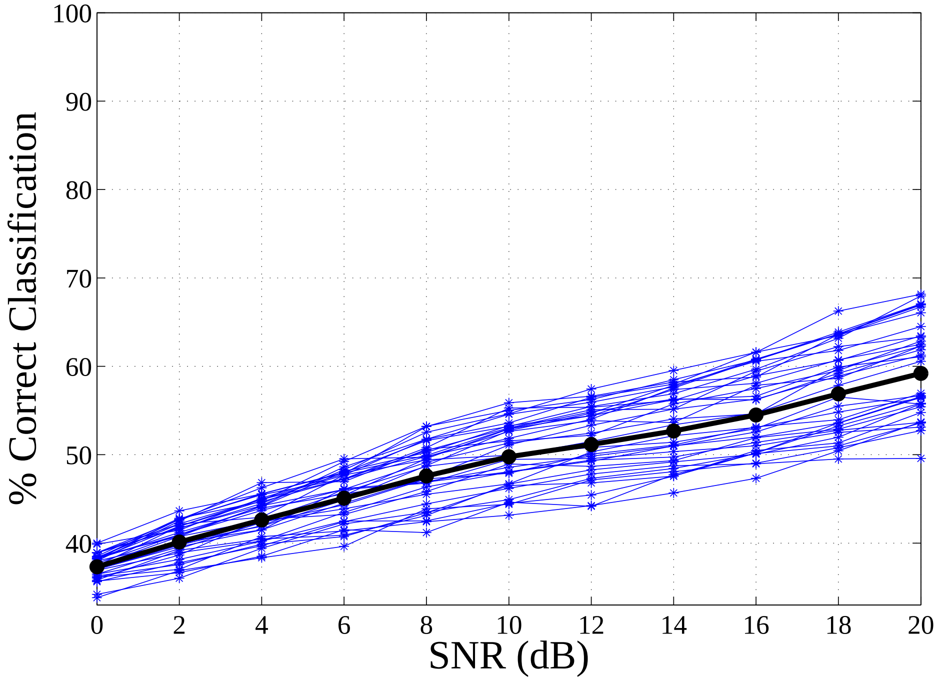


Figure 3.13: 915 MHz CC1000 MDA/ML classification: Quantitatively selected *bottom-ranked* reduced-dimensional (N_F=72 of 144) performance for N_{Prm}=35 3-class permutations. Permutation average shown as filled circle markers.

in Fig. 3.14 using sorted p-values. Components left of the vertical line are top-ranked features sufficient to achieve near-equivalent reduced and full-dimensional performance.

These results establish the desired one-to-one relationship between p-value and component relevance to classification. Through simple inversion, a sorted descending plot of $1/p$ values would reflect most-to-least relevance, much like ranked eigenvalues do in Principal Component Analysis [GWM$^+$02]. This opens the door to a large body of related research on rank-ordered feature selection that remains to be investigated in support of future research aimed at formalizing an optimal feature selection method using pre-classification $1/p$ values.

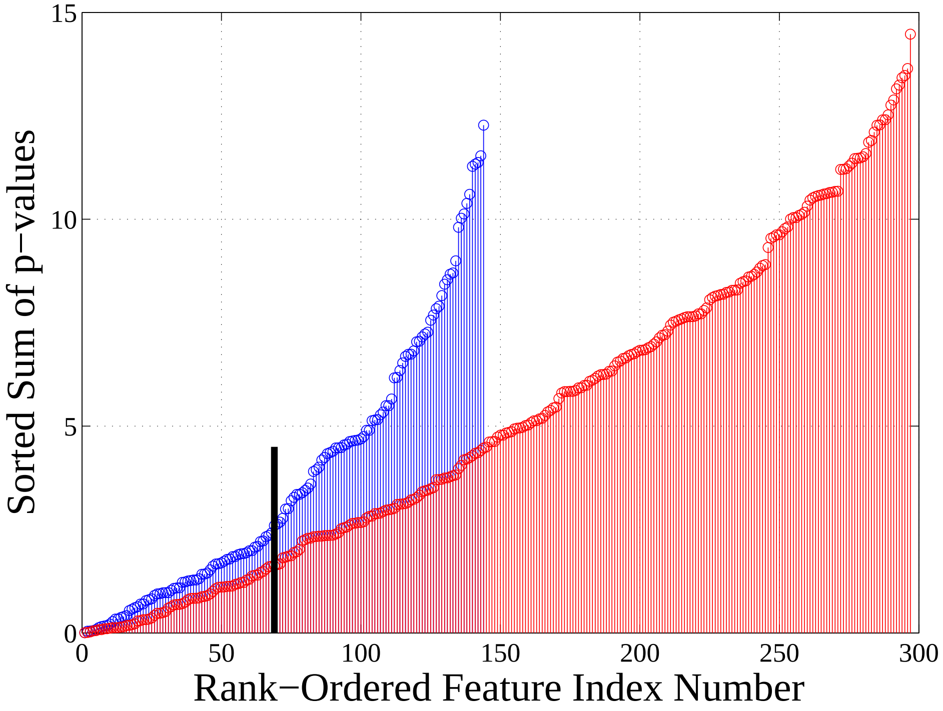


Figure 3.14: Sorted KS-test *p*-value sums for full-dimensional feature sets of 915 MHz CC1000 (N_F=144) and 2.4 GHz CC2420 (N_F=297) devices at *SNR*=8.0 dB. Components left of the vertical line are top-ranked subsets sufficient for near equivalent reduced and full-dimensional performance.

3.7 Summary & Conclusion

The low-cost, low complexity, and low power consumption benefits of ZigBee LR-WPANs make them an attractive alternative for critical infrastructure elements requiring wireless sensing and control. However, the attractiveness is diminished when considering that networks using these devices are relatively easy to exploit using readily available tools. Considering a defense-in-depth approach to mitigating security vulnerabilities, the work here addresses multi-factor PHY-MAC-NWK authentication by adding previously under-exploited PHY information to augment bit-level mechanisms.

PHY features are captured in RF fingerprints and used here to assess device differentiability of like-model 2.4 GHz CC2420 ZigBee transceivers and 915 MHz CC1000 transceivers under varying *SNR* conditions. Performance of a *quantitative*, statistic-based, pre-classification feature selection process is validated and dimensional efficiency demonstrated using an MDA/ML classification process. The contribution here is development and formalization of a concept that was only *qualitatively* assessed in previous related work.

Results here demonstrate that 2.4 GHz CC2420 ZigBee devices can be accurately and reliably discriminated using RF statistical features extracted from signal preamble responses. Most notably, a comparative performance benchmark of $\%C \geq 90\%$ average classification accuracy is achieved for $SNR \geq 8.0$ dB using like-model devices and full-dimensional RF fingerprints (N_F=297 PHY features). Of equal significance, effectiveness of the proposed pre-classification feature selection process is demonstrated using a rank-ordering of KS-test *p*-values, with *p*-value shown to correlate directly with classification feature relevance. The rank-ordering enables reduced dimensional analysis using a small subset (N_F=33) of most relevant features which achieve nearly equivalent *%C* as the full-dimensional (N_F=297) feature set; less than a 1.0 dB trade-off in required *SNR* is required for an approximate 88% reduction in required features.

While not as effective from an overall average *%C* versus *SNR* perspective, classification performance using like-model 915 MHz CC1000 ZigBee devices is promising and the potential for PHY-based multi-factor authentication exists; $\%C \geq 80\%$ performance is achieved for $SNR \geq 8.0$ dB. Of greater importance to continued development of a PHY-MAC-NWK authentication framework, reduced-dimensional results for the CC2420 and CC1000 devices collectively demonstrate the desired one-to-one relationship between *p*-value and feature relevance to classification. Using rank-ordered 1/*p*-values to identify most-to-least relevant features opens the doorway for optimal selection methods similar to what is used in PCA eigenvector analysis.

IV. Wireless Critical Infrastructure Protection using Low-Cost RF Fingerprinting Receivers

4.1 Introduction

One promising solution for securing LR-WPANs without placing additional burdens on end devices is Radio Frequency (RF) fingerprinting. In such systems, an air monitor passively observes all LR-WPAN packets and identifies message spoofing (e.g., packet replay attacks) through device-unique RF fingerprints. Wireless device classification accuracies exceeding 99% have been demonstrated using high-end signal collection receivers (cost exceeding 50K U.S. dollars), including: a 4 Gigasample-per-second (Gsps) oscilloscope [DC09], 8 Gsps oscilloscope [DLCED10], 50 Gsps oscilloscope [PDG11], a 95 Gsps Agilent E3238S signal intercept system [DRT12][RTM12], and an Agilent PSA E4448A Spectrum Analyzer combined with a 4 Gsps oscilloscope [RSC12]. The high cost of these signal receivers prohibits their use in practical RF fingerprinting systems. Thus, techniques developed using high-end receivers must be successfully transitioned to low-cost (less than 2K U.S. dollars) hardware such as the Universal Software Radio Peripheral (USRP). Transient-based fingerprinting requires at least 4 Gsps [DC09][DLCED10], which is not possible on the USRP which is limited to 25 Msps. Spectral fingerprinting using wireless preambles, however, was recently demonstrated with the USRP [RSC12][RSC14]. Initial results suggest lower device differentiation accuracy and higher receiver-specific variability with USRP receivers than with high-end receivers.

The inexpensive analog components in low-end receivers contribute noise and variability during signal reception, and confound the RF fingerprinting process. While some distortion is unavoidable, the hypothesis is that the variability in collection center frequency and environmental noise can be mitigated through post-collection signal processing. Herein, signal processing techniques to mitigate RF fingerprinting limitations of low-cost receivers are demonstrated. RF fingerprinting performance is compared between two RF receivers under identical signal collection conditions, i.e., a high-end National Instruments (NI) PXIe-1085 system and a low-cost NI USRP-2921 were used to simultaneously collect device emissions for a given experimental setup. Accurate device spoofing identification in scenarios involving real-world attack hardware and actual smart utility meters are also demonstrated.

The rest of this chapter is organized as follows: Section 4.2 provides a review of RF fingerprinting. Section 4.3 describes realistic threats to critical infrastructure networks. Section 4.4 details the RF fingerprinting methodology. Section 4.5 presents device classification results for realistic attack hardware. Section 4.6 reports device classification results for smart utility meters. In Section 4.7 accurate device identity verification and anti-spoofing capability using RF fingerprints generated with a low-cost software defined radio are demonstrated. Finally, Section 4.8 presents main conclusions and suggests areas for future work.

4.2 RF Fingerprinting Background

The earliest RF fingerprinting systems were developed by militaries to differentiate among friendly and hostile radar transmissions [HY12]. Costs associated with RF fingerprinting have declined over the last few decades to such a degree that commercial cell phone companies can now use some form of RF fingerprinting to detect device cloning [KS99]. In order to be commercially viable, RF fingerprinting low-cost LR-WPANs in critical infrastructure applications must be practical and use the smallest, least expensive receiver technology.

The authors in [RSC12][RSC14] are among the first to attempt robust RF fingerprinting using low-cost USRP receivers. Their RF fingerprints consist solely of Power Spectral Density (PSD) features of IEEE 802.11a (5GHz WiFi) preambles. IEEE

802.15.4-based LR-WPANs (e.g., ZigBee) likewise feature a preamble at the start of every burst transmission that is amenable to RF fingerprinting. However, recent work with high-end receivers [RTM12] reports that RF fingerprints based solely on PSD features under perform those based on time-domain features. The hypothesis is that RF fingerprinting performance using the USRP can nearly match that of high-end receivers with proper feature selection and a sufficiently robust processing.

Instead of using PSD features, this chapter utilizes a series of instantaneous time-domain features that improve the relative fingerprinting accuracy of the USRP; the robust RF fingerprinting methodology is presented in Section 4.3.

4.3 RF Fingerprinting Methodology

Since the USRP sampling rate is insufficient for transient-based RF fingerprinting [DC09][DLCED10], and recent works highlight PSD-based fingerprinting accuracy limitations [RTM12][RSC12][RSC14], the process outlined in this chapter instead leverages instantaneous time-domain features of the wireless preamble. Robust signal processing techniques including frequency down-conversion and baseband filtering strategies further improve performance.

4.3.1 RF Collection Topologies.

In order to compare the relative RF fingerprinting performance of high-end and low-cost receivers, as many parameters as possible are controlled during signal collection. Fig. 4.1. illustrates the collection topology. Table 4.1 lists the parameters controlled between collections made on the high-end NI PXIe-1085 system and the low-cost NI USRP-2921. Six Atmel RZUSBsticks serve as the fingerprinted transmitters, each transmitting 600 IEEE 802.15.4 packets toward both collection receivers at the same time. RZUSBsticks are selected as transmitters because they are the first hardware supported by KillerBee LR-WPAN attack tools. All previous work on LR-WPAN RF fingerprinting investigated the CC 2420 transceiver, so selection of the RZUSBstick broadens the literature to a new transceiver type (Atmel AT86RF230). During an actual spoofing attack, the malicious device will most likely be transmitting from a different location than the impersonated device and with different hardware than used on the victim LR-WPAN. This variance in location and hardware add to the distinctiveness of an attacker's RF fingerprint. Signal collection described in Table 4.1 and Fig. 4.1. is thus a worst-case scenario as would be experienced from an RF fingerprinting perspective, since all of the transmitters differ in only subtle physical variations in hardware due to manufacturing tolerances.

In a second collection scenario the USRP is used to fingerprint three OpenWay CENTRON Smart Meters at Oak Ridge National Laboratory. This expands the RF fingerprinting literature to yet another LR-WPAN hardware type. The smart meter transmit power significantly exceeded 1 mW, so short-range line-of-sight collection as in Fig. 4.1. was not practical without significant attenuation. In order to collect the smart meter transmissions without saturating the USRP receiver, a -30 dB attenuator was added between the collection antenna and the USRP (Fig. 4.2.). The high-end collection receiver is not portable enough to move to the stationary smart meter test bed, so direct high-end versus low-cost comparisons are only conducted using results from the first collection scenario (Fig. 4.1.).

4.3.2 Signal Collection Methodology.

The signal collection methodology was consistent between the receiver systems: NI PXIe-1085 and NI USRP-2921. Both systems record RF in-phase and quadrature (I/Q) data as 16-bit integers, sampled at 20 Msps. This file format takes the form of an interleaved array

$$[I_0\ Q_0\ I_1\ Q_1\ I_2\ Q_2\ ...\ I_n\ Q_n],$$

Table 4.1: RF collection parameters for high-end NI PXIe-1085 and low-cost NI USRP-2921 receivers.

Parameter	Value
Tx-RX separation distance	2 m
TX-RX height above floor	1 m
Collection time frame	Concurrent
Transmitter	Atmel RZUSBstick
Transmit power	1 mW
Transmitter orientation	Vertical USB port
Receiver antenna	3 dBi gain VERT2450
Receiver antenna orientation	Vertical

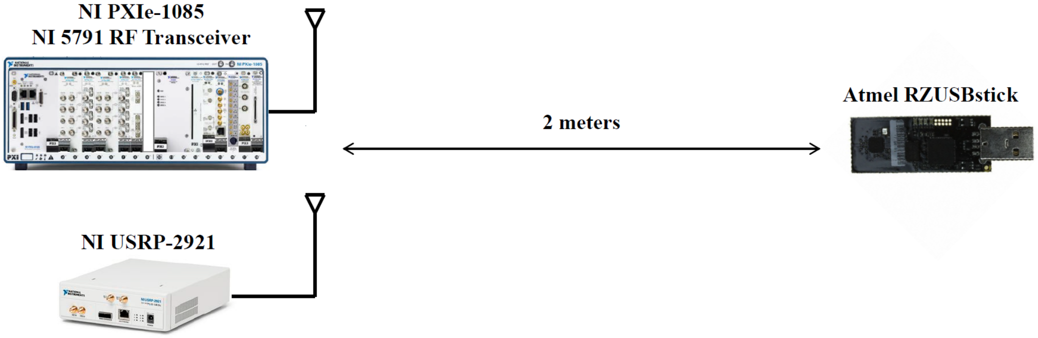


Figure 4.1: RF collection topology for simultaneous collection of RF emissions from six Atmel RZUSBsticks using NI PXIe-1085 and NI USRP-2921 systems.

where n is the number of collected I/Q sample pairs. This interleaved I/Q data is first converted to complex values in the format

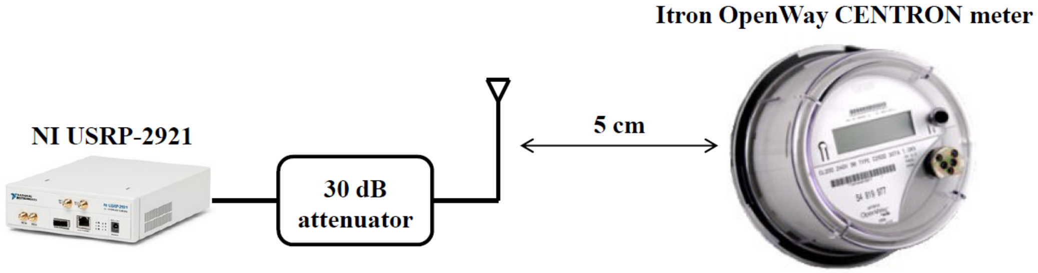


Figure 4.2: RF collection topology for the three OpenWay CENTRON smart meters using the NI USRP-2921.

$$[I_0 + iQ_0,\ I_1 + iQ_1, I_2 + iQ_2, ...I_n + iQ_n],$$

for convenient signal processing in MATLAB. A total of 600 transmission preambles were sampled in this way from six RZUSBsticks using both collection receivers. Transmission detection from background noise was accomplished through amplitude-based leading edge detection using a -6 dB threshold. As outlined in the IEEE 802.15.4 standard, the first 128 μs of each transmission constitutes the preamble. At 20 Msps the first 2560 instantaneous I/Q samples represent the preamble region of each transmission. Fig. 4.3. illustrates a representative IEEE 802.15.4 preamble baseband response, which begins at sample number 500 and ends at sample number 3060. The vertical dashed lines indicate division of the preamble into 32 fingerprint regions, a process further discussed in Section 4.3.3. The transmitter operating frequency was IEEE 802.15.4 channel 26 (2.480 GHz) for all collections to mitigate interference from nearby IEEE 802.11g traffic (2.401-2.473 GHz). Collected signal-to-noise ratio (SNR) was approximately 30 dB on the PXIe-1085 and 24 dB on the USRP.

Inter-device variability in RF fingerprint performance on USRPs was noted in [RSC12]. To mitigate possible variability in collection center frequency due to clock

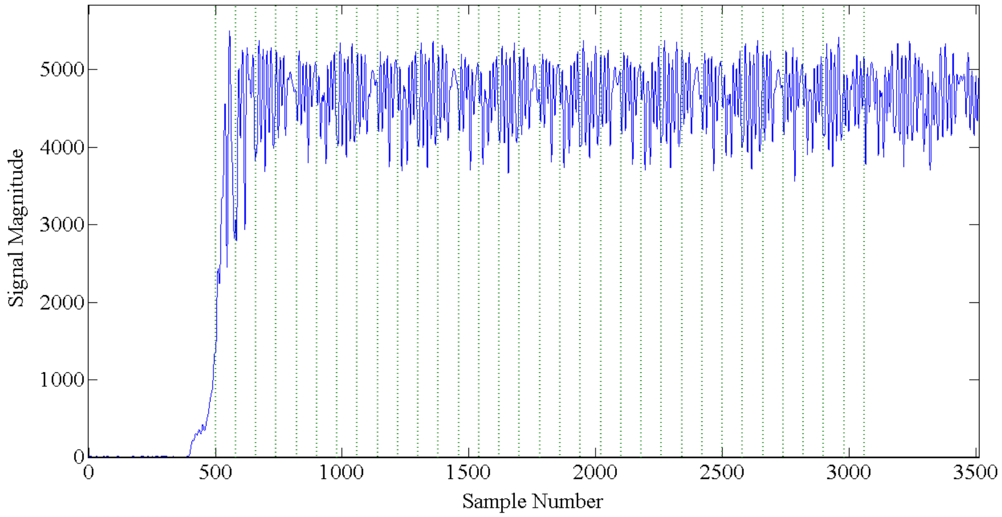


Figure 4.3: Provision of baseband LR-WPAN preamble magnitude response into 32 subregions for RF fingerprinting.

skew, collection center frequency is set to 3 MHz below the transmission center frequency (2.477 GHz versus 2.480 GHz). Fig. 4.4. illustrates normalized Power Spectral Density (PSD) of a representative transmission collected on a USRP using this 3 MHz offset. The 2 MHz-wide spectrum of the transmitter is notably higher than the noise floor and is clearly evident, centered 3 MHz above baseband.

The collected transmission is down-converted to baseband using gradient-based frequency estimation performed using MATLAB. A W_{BB} = 1 MHz-wide, 4th-order Butterworth filter removes background noise from outside the IEEE 802.15.4 channel, resulting in a low-noise, baseband representation of the collected transmission (Fig. 4.5). Background noise filtering was not discussed in [RSC12][RSC14], which may have contributed to the erratic RF fingerprinting performance observations therein.

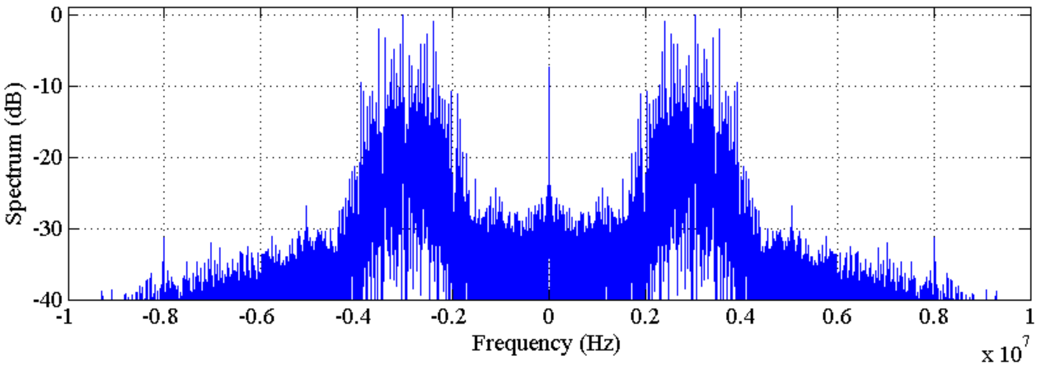


Figure 4.4: Normalized Power Spectral Density (PSD) of an IEEE 802.15.4 transmission collected at 20 Msps using a 3 MHz center frequency offset.

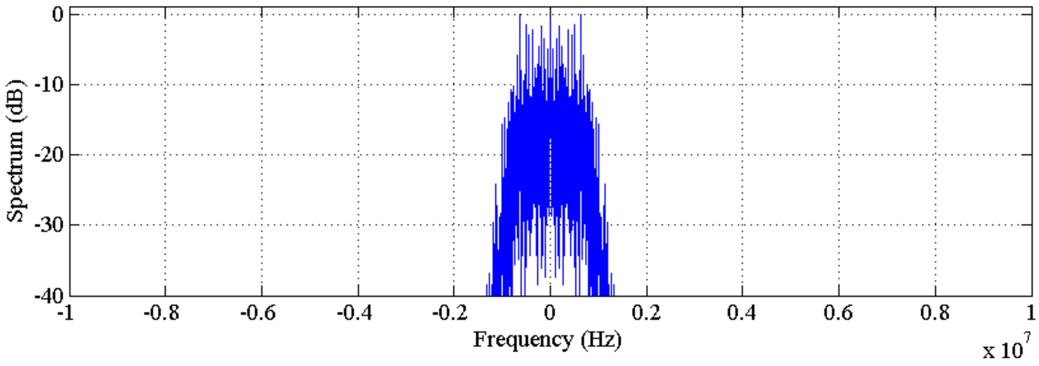


Figure 4.5: Normalized PSD of an IEEE 802.15.4 transmission down-converted to baseband and filtered with a W_{BB}=1 MHz 4th-order Butterworth filter.

4.3.3 RF Fingerprint Generation.

The RF fingerprint (F) for a signal is derived from its instantaneous amplitude (a), phase (ϕ) and/or frequency (f) characteristics. More specifically, the sequences {a[n]}, {ϕ[n]}, and/or {f[n} are generated from complex samples of the signal region of interest, centered (mean removal) and then normalized (division by maximum value) [RTM12]. Instantaneous features are computed from I/Q characteristics of collected preambles.

Consistent with the RF fingerprinting process introduced in [RTM12], 32 preamble sub-regions, or four regions per each of the eight repeated LR-WPAN preamble sub-responses are used (Fig. 4.3) The preamble as a whole serves as the 33rd region.

4.3.4 MDA/ML device classification methodology.

Statistical RF fingerprints are generated using (3.3) for device preamble transmissions from six Atmel RZUSBStick transmitters. The resultant RF fingerprints are classified here using a Multiple Discriminant Analysis, Maximum Likelihood (MDA/ML) process in MATLAB. MDA is a straightforward extension of the Fisher Linear Discriminant process when discrimination of more than two classes (devices) is required. MDA reduces the higher-dimensional input feature space with the goal of maximizing inter-class separation while reducing intra-class spread [HFT09]. For the six-class problems considered here, MDA/ML projects the multidimensional RF fingerprints into a 5-dimensional space. RF fingerprints are classified as being affiliated with one of six possible classes based on Bayesian decision criteria using prior known probabilities, probability densities, and relevant costs associated with making a decision [DHS99]. For all results presented herein the associate costs are assumed equal for all classes.

The MDA/ML models were developed using a K-fold cross-validation training process with K=5 to improve reliability. This value is consistent with literature which suggests values of K=5 and K=10 are sufficient [DHS99]. The best-performing model generated during the training process is subsequently used to generate testing results using a previously unseen collection of input features from each device. Model training is performed using Ntng=300 randomly-selected collected transmissions from each RZUSBstick, and testing is performed on the remaining Ntst=300 collected transmissions that did not take part in model training.

4.4 RZUSBstick device classification performance

This section evaluates the relative performance of the two collection receivers with respect to device classification using RF fingerprints. RZUSBstick devices were selected as transmitters due to their use in LR-WPAN attacks from KillerBee and Api-do.

MDA/ML inter-device classification results were generated for all six RZUSBsticks using the high-end and low-end receivers. Classification experiments incorporated a total of 600 independent transmissions, each beginning with the IEEE 802.15.4 preamble, and 15 Monte Carlo noise realizations per preamble at each test SNR. Model development used only the first 300 preambles, while testing was independently performed using the second 300 preambles. This resulted in (300 test preambles) x (15 noise realizations) = 4500 total classification decisions per device at each test SNR. This large number of trials reduced the 95% confidence intervals to within the vertical extent of the plotted markers. For visual clarity, confidence interval bars are not presented in classification plots.

4.4.1 Full-dimensional RF fingerprints at a 20 Msps sample rate.

Full-dimensional RF fingerprints include features based on all three signal characteristics (a, ϕ, and f), three statistical features (σ^2, γ, and k), and 32+1 preamble regions, for a total RF fingerprint length of $N_F = 3 \times 3 \times 33 = 297$ features. Both collection receiver sample rates were 20 Msps. Fig. 4.6 presents the full-dimensional classification accuracies for six RZUSBsticks using the PXIe-1085 collection receiver, and Fig. 4.7 presents the full-dimensional classification accuracies for the same six RZUSBsticks using the USRP-2921 collection receiver. The solid black lines with circle markers in the figures show the mean classification accuracy for the collections of the six transmitters.

Classification accuracies between the two collection receivers are substantially more consistent than observations reported in [RSC12][RSC14] for IEEE 802.11a devices. The low-end USRP receiver is capable enough to classify all six devices with an average of

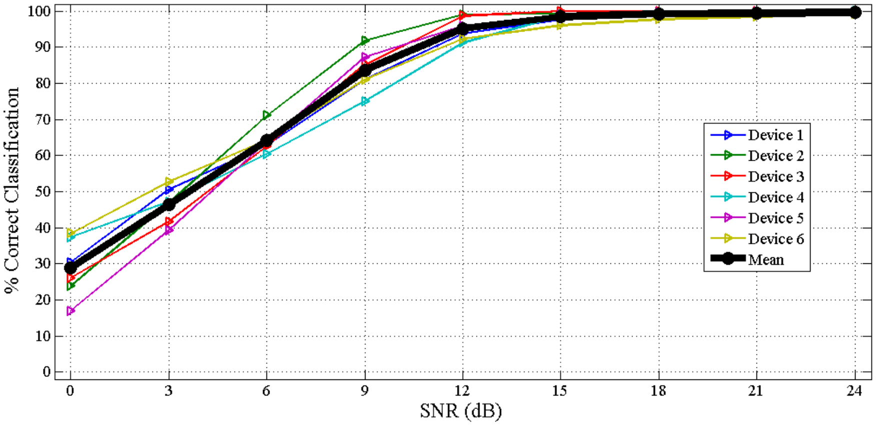


Figure 4.6: Classification accuracy using the NI PXIe-1085 and N_F=297 full-dimensional RF fingerprints at 20 Msps.

90% accuracy when *SNR*≥15 dB. The high-end PXIe-1085 system achieves an average of 90% classification accuracy when $SNR \leq$11 dB. This high-end PXIe-1085 result closely matches findings in [DRT12] using the high-end Agilent E3238S receiver system, where average classification accuracy of seven CC2420 transmitters reached 90% by *SNR*=10 dB. Average device classification accuracy using the USRP is 9% lower than with the high-end PXIe-1085 receiver at *SNR*=12 dB, but this difference narrows to 3% for *SNR*=24 dB. These full-dimensional device classification results are consistent with the intuitive assumption that device classification accuracy using low-cost USRP hardware measurably underperforms that of high-end signal receiver hardware. However, the difference in classification accuracy between low-end and high-end hardware narrows to a few percent under high SNR conditions.

Given the relatively low-cost of software-defined radios such as the USRP, multiple receivers can be purchased for far less than a single high-end receiver. Combining RF fingerprint decisions from multiple low-cost receivers may be an effective strategy to

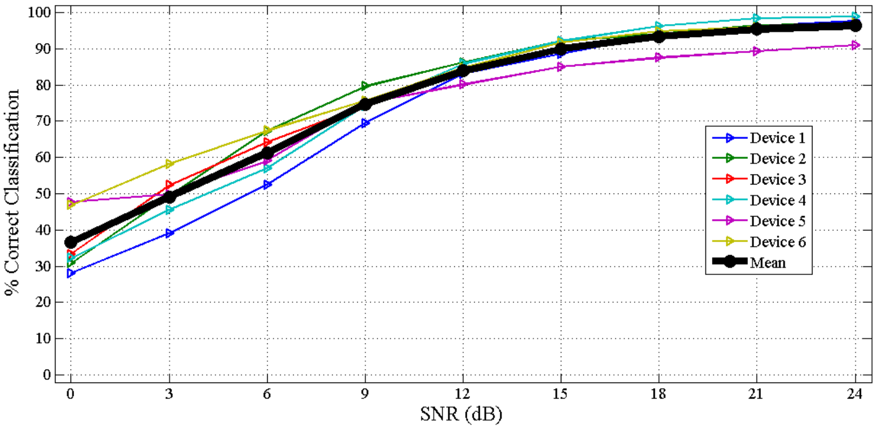


Figure 4.7: Classification accuracy using the NI USRP-2921 and N_F=297 full-dimensional RF fingerprints at 20 Msps.

improve device classification performance. For example, if two out of three USRP-based air monitors determine the same device classification of an incoming packet, their decision may be more accurate than that of the dissenting receiver.

4.4.2 Full-dimensional RF fingerprints at a 5 Msps sample rate.

While the PXIe-1085 and USRP-2921 both support sampling rates as high as 25 Msps, it is not clear that a higher sampling rate necessarily results in greater RF fingerprinting accuracy of LR-WPAN transmitters. This section investigates full-dimensional RF fingerprinting at a reduced sampling rate of 5 Msps. The original RF signal collections were properly decimated from 20 Msps to 5 Msps by utilizing every fourth I/Q sample and excluding the rest from the RF fingerprinting process. A sample rate of 5 Msps is approximately the lowest possible sample rate with which the near-baseband RF collection process will work, since a sample rate of 5 Msps on the USRP equates to a collection bandwidth spanning 2.5 MHz above and below the collection center frequency (IEEE 802.15.4 channel width is 2 MHz). If RF fingerprinting

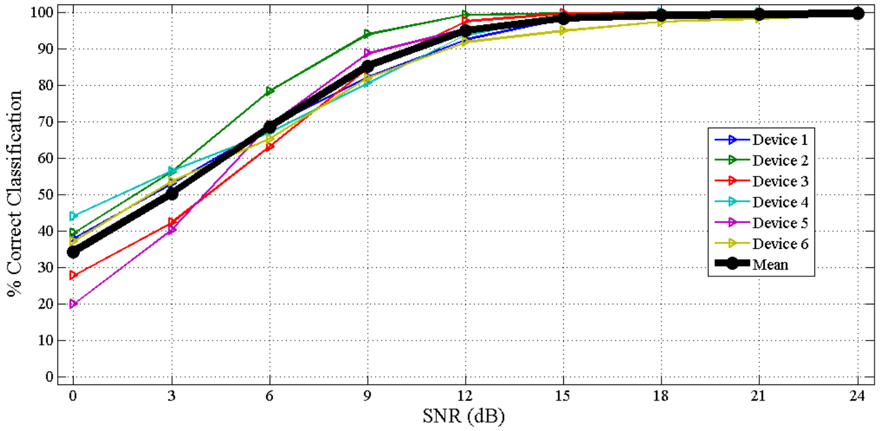


Figure 4.8: Classification accuracy using the NI PXIe-1085 and N_F=297 full-dimensional RF fingerprints at 5 Msps.

at low sample rates is effective, it decreases hardware requirements of RF air monitor hardware deployed in operational systems.

Fig. 4.8 presents the full-dimensional classification accuracies at 5 Msps while using the PXIe-1085 collection receiver, and Fig. 4.9 presents the full-dimensional classification accuracies at 5 Msps while using the USRP-2921 collection receiver. There is a negligible functional difference between device classification accuracies at 20 MHz and 5 MHz while using the high-end PXIe-1085. Similarly, classification accuracies are functionally indistinguishable while using the USRP-2921 at 20 Msps and 5 Msps. A sample rate of 5 Msps meets the Nyquist requirement for 2 MHz IEEE 802.15.4 signals and also appears to provide maximum RF fingerprinting performance. Low-cost RF receiver hardware that supports 5 Msps may be practical systems for RF fingerprinting of LR-WPAN devices.

4.4.3 Phz-only RF fingerprints at a 20 Msps sample rate.

Investigation into *relative relevance* among instantaneous amplitude, phase, and frequency characteristics toward device classification was previously investigated in

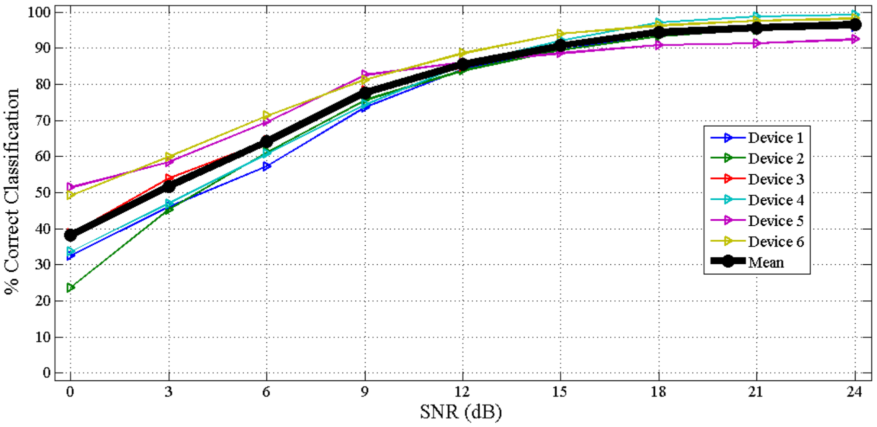


Figure 4.9: Classification accuracy using the NI USRP-2921 and N_F=297 full-dimensional RF fingerprints at 5 Msps.

[RTM12] for the high-end Agilent E3238S system and CC 2420 LR-WPAN transmitters. Results strongly suggested that instantaneous phase features were the most useful for inter-device differentiation, followed by frequency, and that instantaneous amplitudes tended to be the least relevant. Instantaneous phase features were robust enough that phase-only fingerprints (99 features long) were as effective as using full-dimensional fingerprints of 297 features. The advantage of phase-only fingerprints is that they require calculation and processing of only one-third the number of RF features. Fig. 4.10 reports device classification accuracies using RF fingerprints of only 99 instantaneous phase characteristics using the PXIe-1085 receiver. Fig. 4.11 reports device classification accuracies using RF fingerprints of only 99 instantaneous phase characteristics using the USRP-2921 receiver.

Consistent with high-end receiver results in [RTM12], phase-only RF fingerprints are as effective as full-dimensional fingerprints when the high-end PXIe-1085 system serves as collection receiver. Conversely, phase-only classification results while using the USRP-2921 as the collection receiver significantly underperform full-dimensional RF

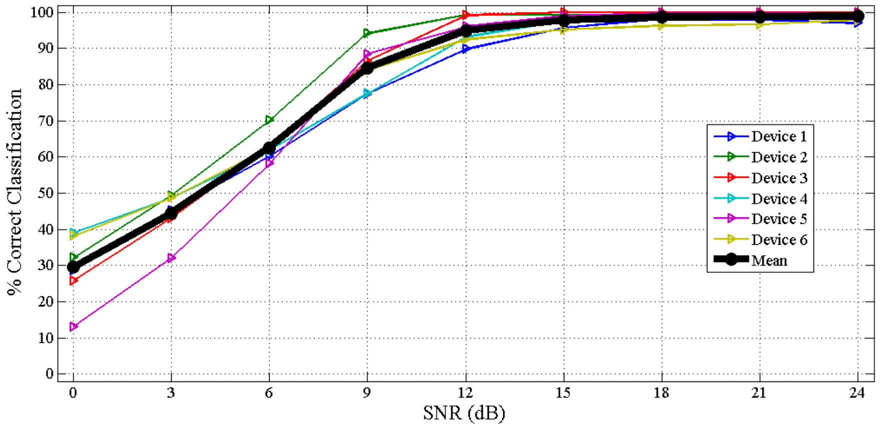


Figure 4.10: Classification accuracy using the NI PXIe-1085 and N_F=99 Phz-Only RF fingerprints at 20 Msps.

fingerprinting. Average device classification using the USRP-2921 falls below 90%, even when SNR = 24 dB. Results indicate that the MDA/ML device fingerprint model incorporates additional RF fingerprint characteristics (more frequency or amplitude traits) when a low-end receiver is used than when a high-end RF receiver is used. The hypothesis is that the inexpensive analog components within the USRP introduce additional RF fingerprint distortion that the MDA/ML model overcomes by diversifying the instantaneous characteristics given the most weight during model development.

To test this hypothesis device classification is performed using both collection receivers and RF fingerprints consisting of only one of the three RF characteristics (amplitude, phase, or frequency). The classification results report relative feature relevance to MDA/ML model development on signals collected on the two receivers (Fig. 4.12 and Fig. 4.13). The phase-only classification results reported in Fig. 4.12 and Fig. 4.13 are the classification means in Fig. 4.10 and Fig. 4.11, respectively.

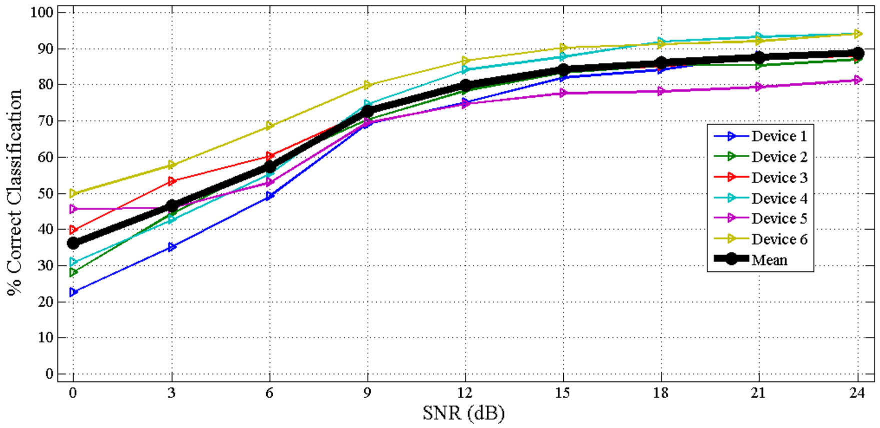


Figure 4.11: Classification accuracy using the NI USRP-2921 and N_F=99 Phz-Only RF fingerprints at 20 Msps.

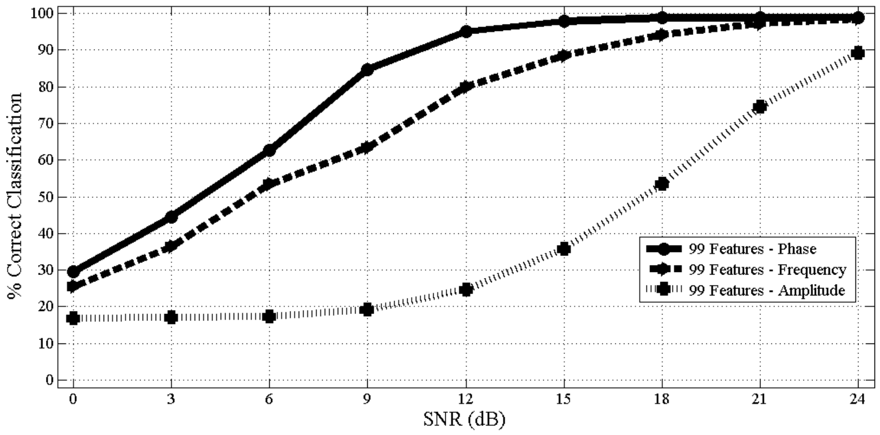


Figure 4.12: Classification accuracy using the NI PXIe-1085 and N_F=99 single-characteristic RF fingerprints at 20 Msps.

While the relative relevance of instantaneous RF characteristics remains the same among the Agilent E3238S [RTM12], PXIe-1085, and USRP-2921 (phase > frequency > amplitude), classification accuracy attainable using any one instantaneous characteristic is

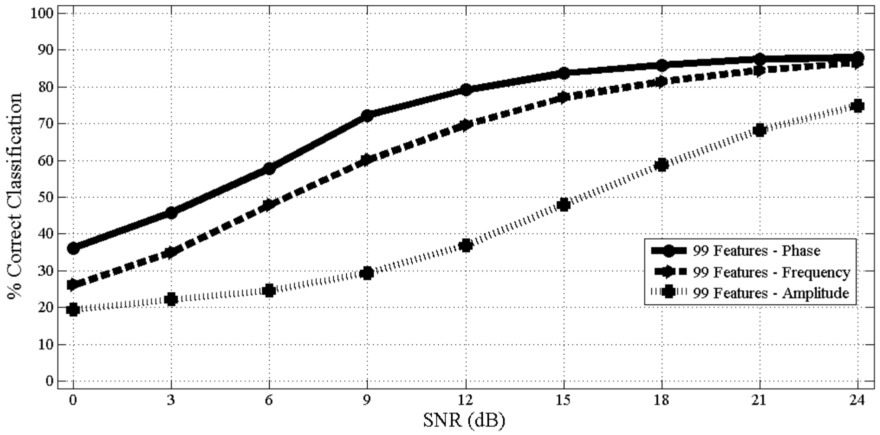


Figure 4.13: Classification accuracy using the NI USRP-2921 and N_F=99 single-characteristic RF fingerprints at 20 Msps.

significantly lower for the USRP-2921; the arbitrary 90% correct classification benchmark is not achieved in any case.

For example, mean device classification accuracy using RF fingerprints consisting solely of phase or frequency characteristics is sufficient for near-100% accuracy for $SNR \geq 21$ on high-end receivers. However, device classification accuracies achieved using equivalent single-characteristic RF fingerprints with USRP-2921 hardware underperform high-end receivers by 10% or greater at SNR = 24 dB. Dimensionality reduction of RF fingerprints generated with low-end collection receivers is still possible through feature ranking, but trivial reduction through selection of a single RF fingerprint characteristic to incorporate is clearly not possible when a low-cost USRP serves as the collection receiver.

4.5 Smart meter classification performance

Smart meter device classification was evaluated using RF fingerprints and the USRP-2921 collection receiver. The high-end PXI-1085 was not portable enough to be relocated to the stationary test bed of OpenWay CENTRON Smart Meters at Oak Ridge

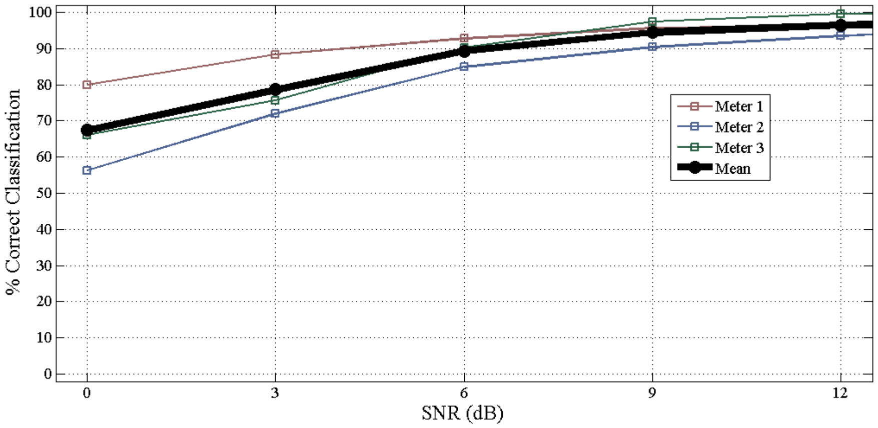


Figure 4.14: Smart meter classification accuracy using the NI USRP-2921 and N_F=297 full-dimensional RF fingerprints at 20 Msps.

National Laboratory. The collection topology is illustrated in Fig. 4.2. The -30 dB signal attenuator placed between the receiver antenna and USRP-2921 decreased the collected SNR to 14 dB. Device classification results are reported for *SNR* in [0, 12] dB since added white Gaussian noise (AWGN) can only be added to the original signal collections to produce the test SNR environments, not further reduced.

Fig. 4.14 presents device classification results using the USRP-2921 and full-dimensional RF fingerprints at 20 Msps. Mean device classification accuracy reaches 90% when SNR = 6 dB and increases to 96% at SNR = 12. Fig. 4.15 presents device classification results using the USRP-2921 and full-dimensional RF fingerprints calculated from properly decimated RF collections at 5 Msps. As observed earlier for the six RZUSBsticks, mean device classification accuracy does not diminish when the sample rate decreases from 20 Msps to 5 Msps. This is additional evidence that 5 Msps is sufficient for maximum RF fingerprinting performance involving IEEE 802.15.4 devices.

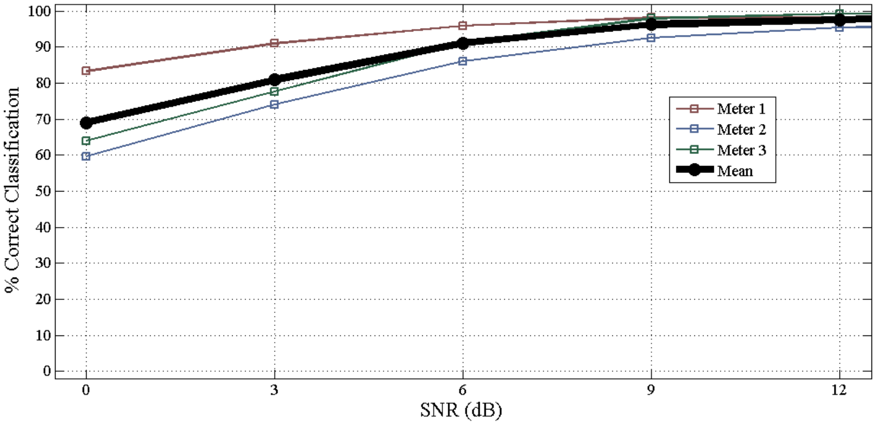


Figure 4.15: Smart meter classification accuracy using the NI USRP-2921 and N_F=297 full-dimensional RF fingerprints at 5 Msps

4.6 Device identity verification

Device classification results in earlier sections establish that this RF fingerprinting methodology is effective at inter-device differentiation. In this section, the capability is demonstrated for RF fingerprinting to detect device spoofing attacks against critical infrastructure LR-WPANs.

As described previously, RF fingerprints are generated for nine devices: six RZUSBsticks and three smart meters using the USRP-2921 as the collection receiver. In an RF-fingerprint-defended LR-WPAN, the air monitor system trains on RF fingerprints calculated from transmissions made by its trusted member devices. Spoofing attacks originate from untrusted devices with hardware-unique RF fingerprints that do not exactly match those of any trusted device in the LR-WPAN.

4.6.1 Device verification scenario.

The verification methodology in [DRT12] was adopted here with a subset of authorized devices used for air monitor training, and the remaining devices used to perform spoofing attacks against each of the authorized devices. Since RZUSBsticks are a popular LR-WPAN hardware attack platform, three RZUSBsticks (Devices 1-3, as labeled in Figures 4.6-4.11) served as the spoofing devices. The three remaining RZUSBsticks (Devices 4-6) and three smart meters (Meters 1-3) formed the pool of authorized LR-WPAN devices. The combination of smart meters and RZUSBsticks is consistend with a smart grid LR-WPAN implementation using interconnected smart meters and industrial appliances.

First, an MDA/ML device classification model was generated for the six authorized devices using N_{TNG}=300 preamble-based full-dimensional RF fingerprints each, as described in Sections 4.5-4.6. This created a five-dimensional Fisher projection maximizing inter-device differentiability. RF fingerprints from the three spoofing devices underwent the same Fisher projection as the authorized devices. Each of the three spoofing devices was introduced as an impersonator for each of the six authorized devices, for a total of $3 \times 6 = 18$ spoofing scenarios. SNR was introduced for verification and spoofing rejection was assessed at *SNR*=12 dB.

The posterior output variable from MATLABs classify function provides the verification test statistic for a spoofing device as it impersonates each authorized device. Spoofing device verification is assessed by inputting the posterior output into MATLABs ROC (Receiver Operating Characteristic) function, which yields verification performance curve data.

4.6.2 Device verification accuracy.

Given spoofing Device *j* presenting a claimed identity of Device *i*, there are two probabilities used to generate verification performance curves for spoofing scenarios: 1) P[Di|Fi] provides a measure of how much authorized Device *i* projected fingerprints “look like” authorized Device *i*, and 2) P[Di|Fj] provides a measure of how much spoofing Device *j* “looks like” authorized Device *i*. These probabilities were used to generate results presented in Fig. 4.16. The ROC legend for each of the 18 spoofing scenarios is in the format {spoofing device: spoofed device}.

Results in Fig. 4.16 are interpreted as follows. The vertical axis represents the probability that the authorized device is recognized as legitimate (True Verification Rate), and thus accepted by the LR-WPAN. The horizontal axis represents the probability that the spoofing device successfully impersonates the authorized device (Rogue Accept Rate). A ROC curve that approaches the upper left corner of Fig. 4.16 (TVR=100% and RAR=0%) indicates that a statistical threshold exists where all transmissions from an authorized device are accepted as legitimate and all spoofing attacks are rejected. Conversely, ROC curves removed from the upper left corner indicate only imperfect spoofing detection.

In 16 out of 18 spoofing scenarios (collection of curves near the upper left corner), a threshold of TVR>90% authorized packet acceptance resulted in less an RAR<2% acceptance of spoofed packets. Spoofing detection was 100% accurate for all scenarios in which Dev 1-3 impersonated the three smart meters. In the two most challenging scenarios (Dev 1: Dev 4 and Dev 3: Dev 6) the TVR>90% threshold resulted in spoofed packet acceptances of 54% and 32%, respectively. It is important to recall that the RZUSBsticks were fingerprinted under atypically challenging conditions where many factors were controlled that would otherwise have contributed to inter-device differentiability, including device position and transmit antenna orientation. Given the challenges imposed, the robustness of this verification process is evident. Even LR-WPAN devices of the same hardware type and in the same antenna orientation can be reliably verified using RF fingerprints generated from signals collected on a low-cost USRP receiver.

A useful technique for visually representing RF fingerprints is through RF-DNA (Radio Frequency - Distinct Native Attributes) markers, adopted here from [DRT12]. Fig. 4.17 illustrates average RF-DNA responses for the six RZUSBstick transmitters, generated using the NI USRP-2921 at 20 Msps. Averages were calculated based on 400 preambles at

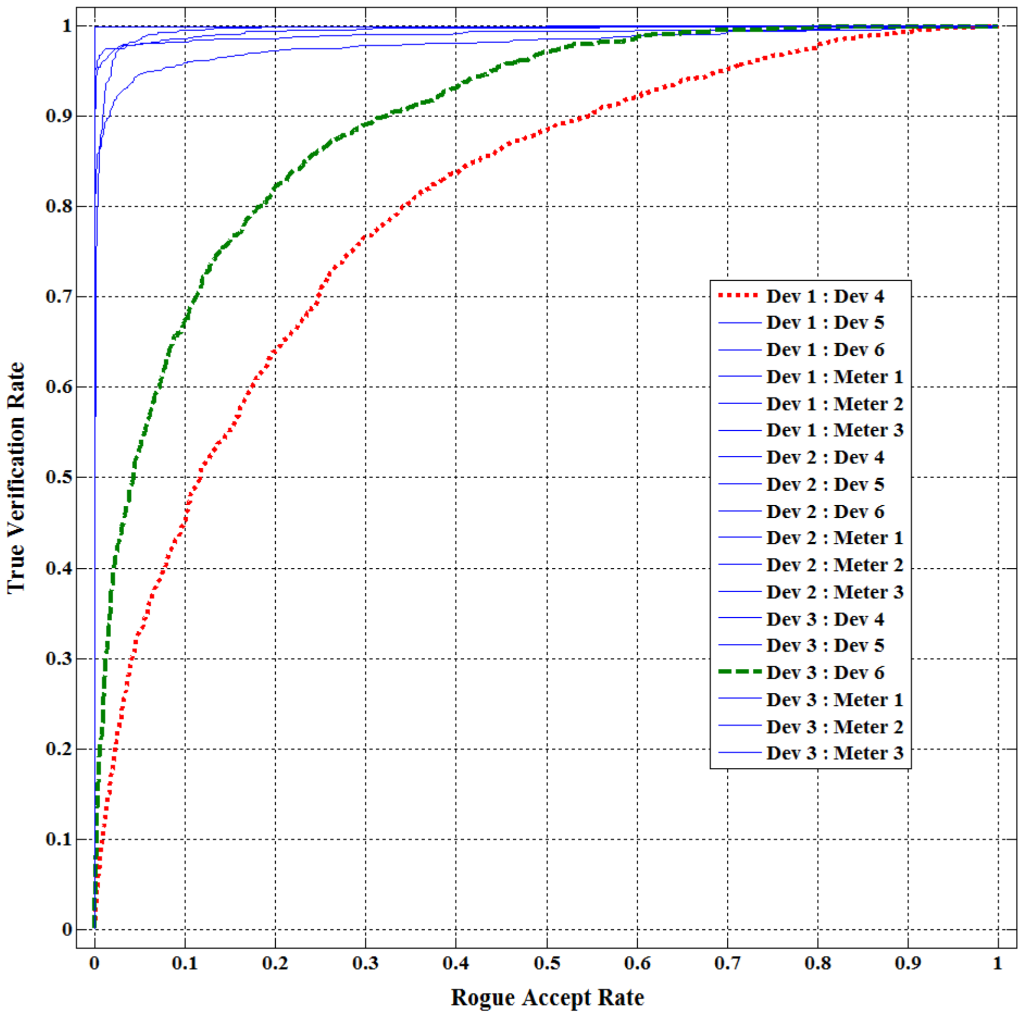


Figure 4.16: Device verification using the USRP-2921 and full-dimensional RF fingerprints at SNR=12 dB.

SNR=12 dB. Full-dimensional NF=297 RF fingerprints include 99 markers for each of the three statistics (variance (var), skewness (skw), and kurtosis (kur)), as described in Section 4.3. It is important to note that this normalized (within statistic) representation was developed to help visualize feature variation across devices. This particular normalization is not included when using RF fingerprints for classification and verification.

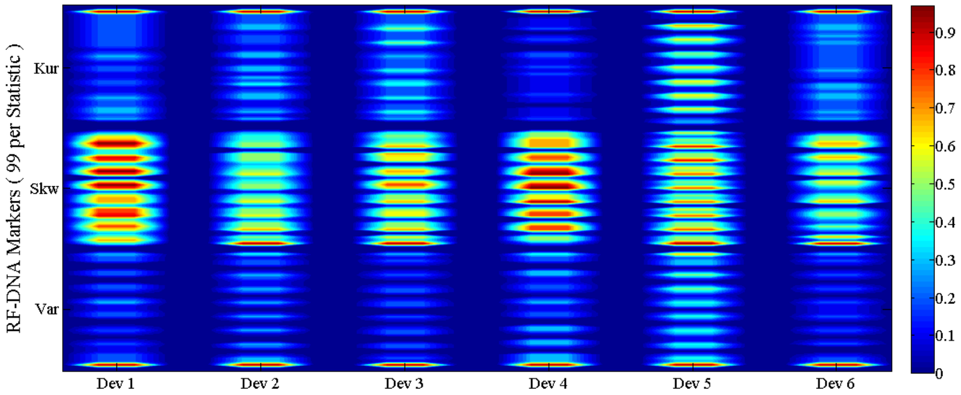


Figure 4.17: Average RF-DNA markers for Devices 1-6 at SNR=12 dB.

Dev 1 and Dev 4 appear the most similar in the RF-DNA visualization shown in Fig. 4.17. The similarity between these devices mirrors the spoofing detection challenge reported by the {Dev 1: Dev 4} ROC curve in Fig. 4.16. The second most challenging spoofing scenario in Fig. 4.16 is the impersonation of Dev 6 by Dev 3. Inter-device similarities are apparent from the RF-DNA markers of Dev 3 and Dev 6 shown in Fig. 4.17.

4.7 Conclusions and Future Work

This chapter demonstrates that reliable RF fingerprinting of critical IEEE 802.15.4 networks is practical using low-cost signal receivers. Distortions introduced by inexpensive analog components are mitigated by conducting signal collection with a small frequency offset and by filtering out background noise effects. Findings suggest that the 25 Msps sample rate of the NI USRP-2921 is not essential to IEEE 802.15.4 fingerprinting, and that lower cost receivers supporting 5 Msps could be sufficient to defend operational systems. It is also found that frequency-based RF fingerprint features are more relevant to device classification when a low-cost signal receiver is used than in high-end receivers. This work represents a significant step toward realizing a practical, low-cost RF fingerprinting solution.

Near-term work continues to lower costs associated with implementing practical RF fingerprinting solutions. The Nuand bladeRF software-defined radio is a fraction of the cost of the NI USRP-2921 and is a promising candidate for upcoming work. IEEE 802.15.4 LR-WPANs are of particular interest in future RF fingerprinting experiments due to their widespread use in critical infrastructure (CI) and supervisory control and data acquistion (SCADA) applications and the unique security challenges they pose.

V. Tuning KillerBee for Critical Infrastructure Warwalking

5.1 Introduction

The National Institute of Standards and Technology highlights the open-source *KillerBee* framework as an important vulnerability research tool for examining critical infrastructure, particularly with regard to ZigBee and smart grids [NIS10]. KillerBee has been under active development since its initial release by Joshua Wright in 2010 [Wri09]. The Api-do project [GBM$^+$12] recently extended KillerBee code to analyze and jam smart meter traffic. Apa and Penagos used KillerBee code during their wireless compromise of industrial facilities over a distance of 60 kilometers [AHP13]. For penetration testers exploring on foot ("warwalking"), the KillerBee tool `zbfind` estimates distance to nearby ZigBee transmitters via received signal strength measurements. Once located, ZigBee devices can be inspected, tampered with, or stolen. Encryption key extraction is a serious concern for low-cost wireless devices [Goo09].

The first empirical evaluation of `zbfind` [RMW12] reveals that its log-distance path loss model, as originally implemented, is inaccurate. In `zbfind` version 1.0 the maximum estimated distance to indoor transmitters is a scant 13 m. Improvements to the `zbfind` distance estimation model (revision r47) incorporate findings from [RMW12], increasing the maximum indoor distance estimate to a more realistic 20.4 m. Empirical evaluation of warwalking tools requires real-world measurements which are challenging and time consuming to collect. However, mathematical modeling alone is insufficient and any model must be operationally validated before it can be relied upon.

The initial KillerBee release supported a single transceiver board, the Atmel RZUSBstick. KillerBee has since expanded its hardware support to include the CC2420 transceiver found in open source hardware such as TelosB motes and the newly-developed ApiMote. To date, these CC2420-based transceiver boards have not been empirically evaluated for use with zbfind. This work evaluates CC2420-based transceiver boards for zbfind-powered warwalking for the first time.

A limitation highlighted in [RMW12] is that the warwalking paths examined are all within office buildings. Alternatively, this chapter extends the literature by investigating warwalks against two targets commonly recognized as critical infrastructure: hospitals and smart utility meters.

This work is organized as follows. Section 5.2 provides background information on distance estimation based on Received Signal Strength Indication (RSSI). Section 5.3 evaluates the suitability of various transceiver boards for zbfind warwalking. Section 5.4 examines distance estimation model accuracy in hospital environments. Section 5.5 examines distance estimation model accuracy outdoors against smart utility meters. Section 5.6 concludes.

5.2 Background

Received radio frequency signal strength declines with increasing distance from the transmitter. Wireless devices quantify received signal strength as RSSI, and conversion from RSSI to dBm is unique to a given hardware configuration. This principle has long been used to approximate transmitter distance because it requires no additional measurements or infrastructure [HLK+10]. Recent works investigate and improve wireless network RSSI-based algorithms in security, tracking, and communication applications [MKP+12][BD13]. Accurate RSSI-based positioning techniques frequently rely upon a grid of static sensor nodes with well-known received signal strength characteristics [TC13]. However, a penetration tester interested in rapidly locating ZigBee transmitters does not have access to such extensive infrastructure. Instead, she must rely solely upon real-time RSSI measurements while walking the target environment.

A popular distance estimation model with demonstrated success indoors is the *log-distance path loss model*. zbfind uses the log-distance estimate

$$d \approx 10^{(A-r)/10P} \tag{5.1}$$

where d is the estimated distance to the transmitter in meters, P is the environmental path loss constant, A is the reference received signal strength at d = 1 m, and r is the received signal strength converted to dBm as measured at an unknown distance from the transmitter.

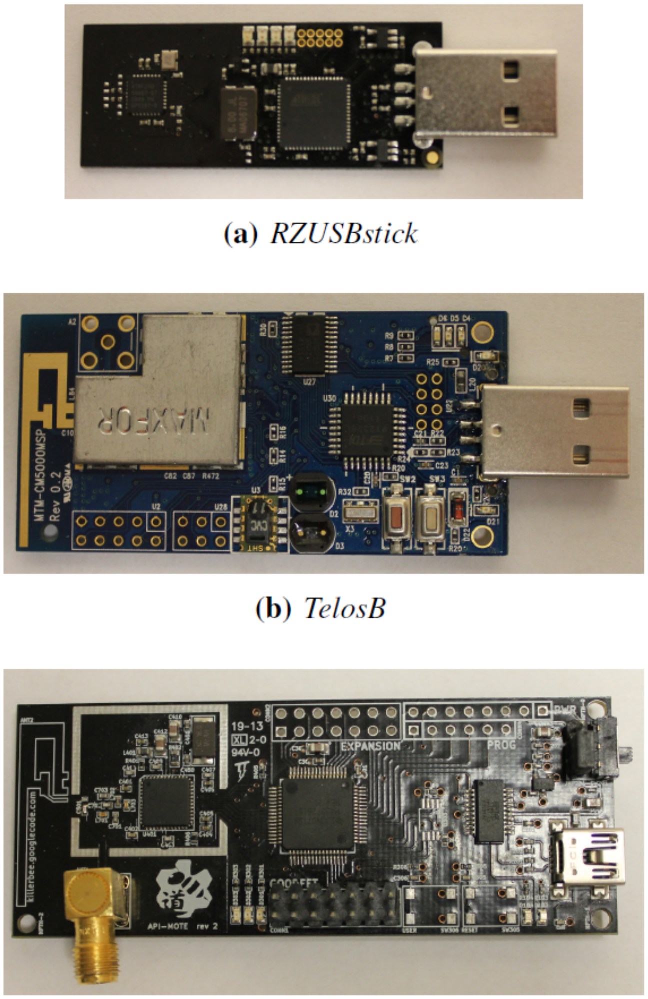
The Atmel RZUSBstick (Fig. 5.1a) features an AT86RF230 transceiver and printed circuit board antenna. The AT86RF230 quantifies RSSI from detected signal energy and stores it as a discrete integer RSSI$\in \{0, 1, ..., 28\}$ in the least significant five bits of its `PHY_RSSI` register [Atm09]. For this device RSSI measurements convert to dBm by

$$r = 3(RSSI) - 91 \tag{5.2}$$

for the computation of A and r in (5.1). An RSSI of zero represents less than -91 dBm received signal power. The RZUSBstick includes a fixed 100 Ω printed 5 dBi gain loop antenna connected directly to the transceiver. RSSI is quantified in steps of 3 dBm, as shown in (5.2). RSSI tolerance is ±5 dBm, so measurements taken under nearly identical reception conditions regularly vary among two or three concurrent RSSI values.

The relationship between RSSI and received signal power on CC2420-based boards is more challenging to establish a priori, since the hardware front ends vary among CC2420-based boards, including the TelosB platform (Fig. 5.1b) and the ApiMote (Fig. 5.1c). The CC2420 conversion equation

$$P = RSSI\_VAL + RSSI\_OFFSET \tag{5.3}$$



(a) *RZUSBstick*

(b) *TelosB*

(c) *ApiMote*

Figure 5.1: Three wireless transceiver boards under test [RMLS14]. Devices shown to scale.

includes an $RSSI\_OFFSET$ term that must be found empirically during system development; the CC2420 datasheet [Ins14] estimates a typical offset of −45 dBm. CC2420 transceivers store $RSSI\_VAL$ as the least significant eights bits of its register 0x13. RSSI tolerance on the CC2420 (±6 dBm) is wider than on the AT86RF230 (±5 dBm).

KillerBee currently supports CC2420-based hardware for packet sniffing, packet injection, and all other bit-layer security evaluation functions originally supported on RZUSBstick hardware. The recently-developed ApiMote is based on the Berkley TelosB mote, but with increased emphasis on hardware modification and security research. Section 5.3 addresses the open question of whether or not these two CC2420-based boards are also compatible with `zbfind` warwalking.

5.3 Preliminary Transceiver Evaluation

This section examines the results of a pilot study to determine the applicability of three transceiver configurations for use with KillerBee's `zbfind` warwalking tool. The three configurations under test are as follows: Atmel RZUSBstick (Fig. 5.1a), TelosB (Fig. 5.1b), and ApiMote (Fig. 5.1c) with an attached 5 dBi dipole antenna oriented vertically.

A 1 mW IEEE 802.15.4 transmitter (Freescale 1321x Sensor Reference Board) is placed at the end of an office corridor 2.5 m wide at its narrowest, represented by the star in Fig. 5.2. Given that a corridor width of 2.5 m is standard in U.S. hospitals, this pilot study serves as a reasonable approximation of a hospital environment prior to trials in real-world hospitals discussed in Section 5.4. A 1 mW transmitter is selected because 1 mW is the nominal transmit power of indoor systems, including TelosB motes and ZigBee-based medical sensors. The transceiver board configurations are connected via USB port to a Dell Precision M4500 laptop computer running Ubuntu 13.10 and KillerBee software, one at a time. Received signal strengths are recorded to a text file for post-collection statistical analysis. Three hundred RSSI measurements at one per second are recorded from each of the three board configurations at ten warwalking distances. Measurements occur at $d \in \{1\ 4\ 7\ ...\ 28\}$ m, for a total of 3000 RSSI measurements per transceiver configuration.

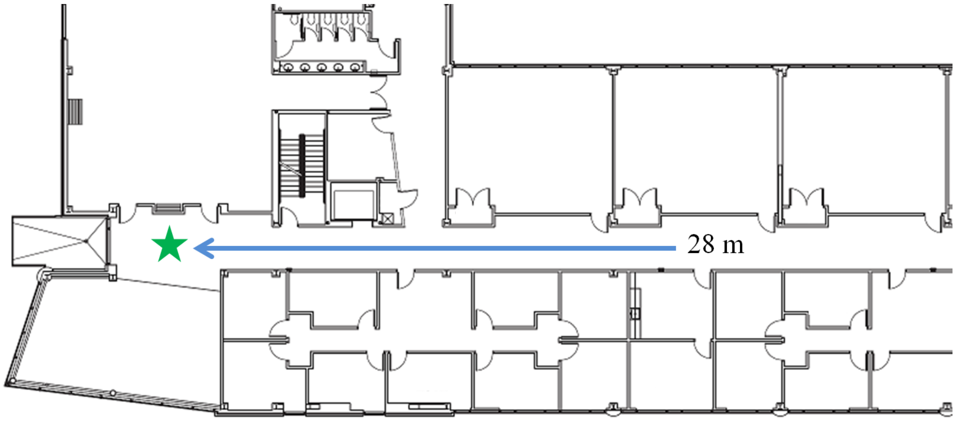


Figure 5.2: Warwalking path during preliminary transceiver evaluation [RMLS14].

`zbfind` provides the user with a distance estimate based solely on RSSI measurements. In order for RSSI to be an effective predictor of distance to the transmitter, there must be a significant correlation between RSSI values and actual distances. Fig. 5.3 plots the relationship between RSSI and distance for the transceiver board configurations under test, and the red dots illustrate all observed RSSI values at each distance. The blue lines show the simple linear regressions.

Distance and RSSI are significantly correlated for RZUSBstick hardware (Fig. 5.3), consistent with [RMW12]. The linear fit has an adjusted coefficient of determination of $R^2_{adj} = 58.2\%$. RSSI observations tend to decrease with respect to distance, particularly for $d \leq 10$ m, where there is a statistically significant ($\alpha = 0.05$) and monotonic decrease in mean observed RSSI from $d = 1$ m (14.82) to $d = 10$ m (6.75). These results suggest that the RZUSBstick is a viable transceiver board for `zbfind` warwalking.

The relationship between RSSI and distance is less significant for the TelosB (Fig. 5.4). The linear fit has an adjusted coefficient of determination $R^2_{adj} = 17.3\%$. While the statistically significant monotonic decrease in mean observed RSSI also holds true for

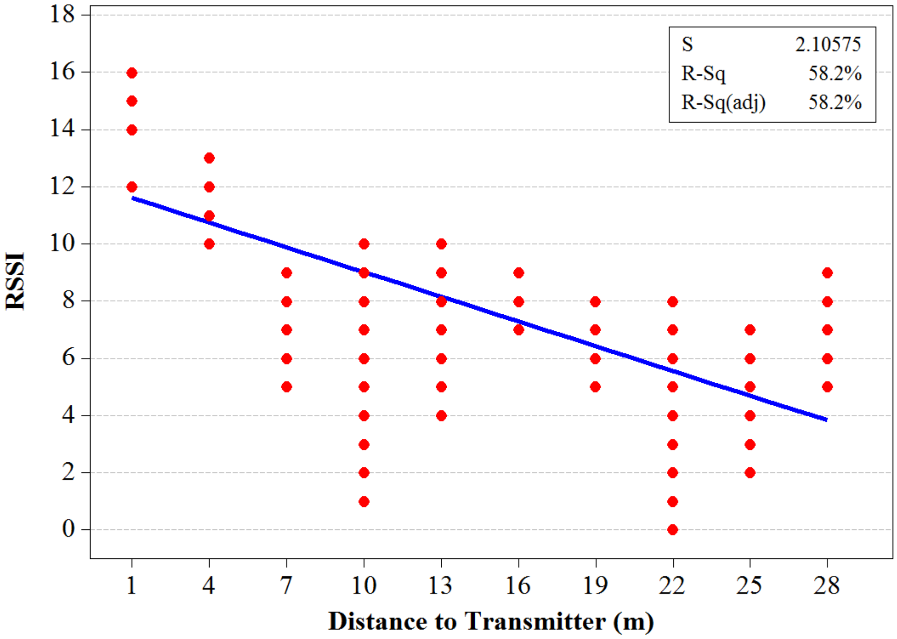


Figure 5.3: RSSI versus distance for the Atmel RZUSBstick [RMLS14].

$d \leq 10$ m, the difference is approximately equal to one (77.95 vs 76.88). This difference in mean RSSI is quantifiable as a real number in positioning systems consisting of a grid of receivers [WLY+14], but is too subtle when RSSI is reported as integer values from a single receiver, as with `zbfind`. Given that ZigBee transmission rates of 1 packet per second or fewer are prevalent [RMSB13], a warwalker would need to pause for a sufficient number of RSSI measurements to accumulate before distance could be estimated. Such measurement accumulation alternatives may warrant exploration in future work, but for the purposes of this work the TelosB is to be dismissed as an impractical for use with `zbfind`.

The ApiMote, with a radio frequency front end based closely off of the TelosB, likewise exhibits a low RSSI/distance correlation. The R^2_{adj} value for the ApiMote with an antenna (Fig. 5.5) is less than 1%. An additional trial with the ApiMote with no antenna on its RP SMA connector (figure omitted for brevity) was also unsuccessful ($R^2_{adj} = 0$).

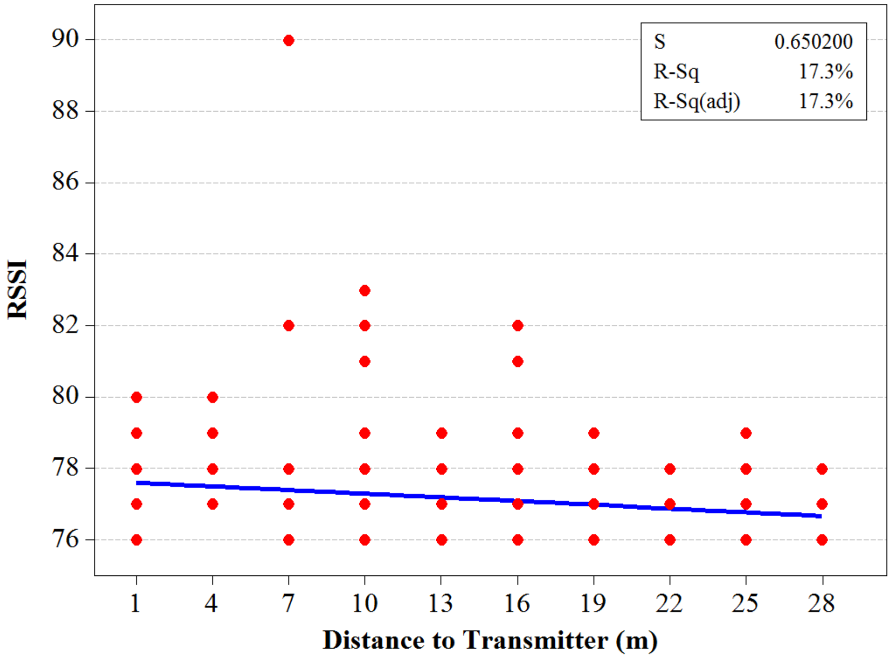


Figure 5.4: RSSI versus distance for the TelosB mote [RMLS14].

Intuitively, the addition of an antenna increases the maximum RSSI values reported, but mean RSSI variability remains low. As with the TelosB, the ApiMote is determined to be impractical for use with `zbfind` as currently designed.

Based on results in this section, subsequent experiments in this work examine Atmel RZUSBstick hardware exclusively.

5.4 Distance Estimation in Hospitals

ZigBee networks in hospitals enable diverse services, from device tracking and distributed sensor measurements to lighting automation. Given the mobile and distributed nature of these networks, establishing physical security is a challenge. In some medical systems this physical security is essentially limited to a notice on the router's plastic enclosure that reads "*Do not remove*." ZigBee devices may be hidden behind potted plants

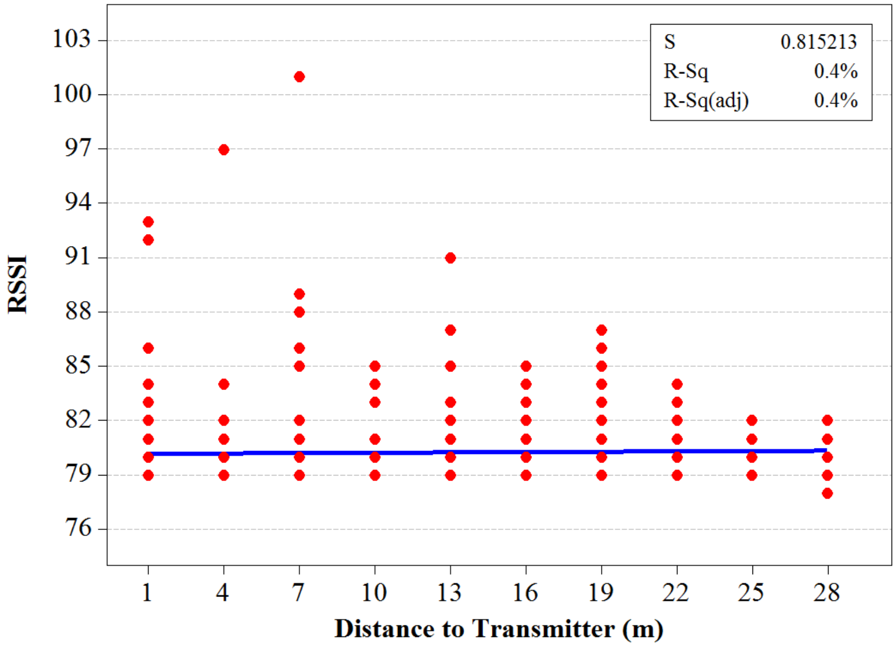


Figure 5.5: RSSI versus distance for the ApiMote with a 5 dBi dipole antenna [RMLS14].

or other obstructions to minimize visibility, so effective warwalking tools are necessary to rapidly locate both benign and malicious transmitters.

There are three broad classes of hospital environments where ZigBee devices may be: 1) corridors, 2) rooms connecting to said corridors, and 3) wide open spaces such as patient waiting areas. Fig. 5.6 illustrates a real-world medical equipment position plot within a military medical facility [Geo10] where devices are in each of these three environments. A warwalker cannot know in advance in which environment the transmitter of interest is located, so an effective distance estimation model must be successful in all three scenarios.

5.4.1 Methodology.

The RSSI measurement methodology herein is closely based on that in [RMW12], but with improved data resolution (3 m versus 5 m), additional measurements at each

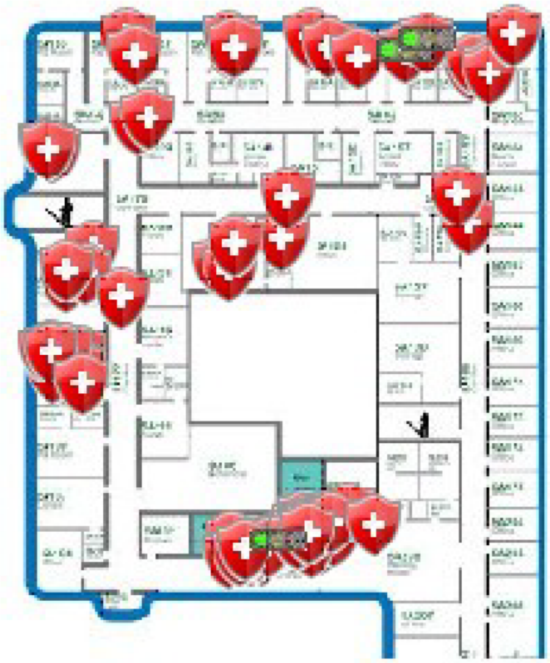


Figure 5.6: Sample equipment location plot in a military hospital [RMLS14].

distance (150 versus 100), and longer maximum distance examined (31 m versus 26 m). A 1 mW Freescale 3121x Sensor Reference Board is placed within a military hospital in each of the three environment classes listed above. RSSI measurements are recorded at three meter warwalking increments from d = 1 m to d = 31 m to the three transmitter locations. One hundred and fifty RSSI measurements are recorded to a text file at each distance for post-collection processing. This number of RSSI measurements allows for k-fold cross-validation with k = 5. Model development incorporates 120 measurements and model testing incorporates the remaining 30 measurements during each of the five rounds to limit model overfitting. Model accuracy is also tested against a warwalk conducted around a real-world ZigBee sensor network operating in a civilian hospital. RSSI measurements from the civilian hospital are not incorporated into model development so that they can be used in unbiased error testing.

Comparative distance estimation model accuracies are compared using mean absolute percentage error (MAPE). For a given distance estimation model, MAPE M is the percentage error defined by

$$M = \frac{100\%}{n} \sum_{t=1}^{n} \left(\left| \frac{A_t - F_t}{A_t} \right| \right) \tag{5.4}$$

where n is the number of fitted points, A_t is the measured value, and F_t is the model fitted value.

5.4.2 RSSI Measurements.

Fig. 5.7 presents mean RSSI observations at each warwalking distance toward three different transmitter locations in a military hospital. The solid black line reports the combined mean for all three environment classes. A trend toward lower RSSI with increasing distance is observed as expected, while the respective plots also reveal the variable nature of RSSI measurements. The solid black line in Fig. 5.9 presents mean RSSI observations in the civilian hospital corridor, restricted to a maximum d = 28 m by the corridor length. The solid black line in Fig. 5.10 reports mean RSSI observations in an office building presented in prior work [RMW12].

5.4.3 Indoor Log-Distance Path Loss Model Evaluation.

Two components of the log-distance path loss model (5.1) require estimation: reference RSSI A and path loss constant P. Cross-validation with k = 5 of the three military hospital warwalks at d = 1 m reports mean RSSI of 13.54 (A = -50.38 dBm), slightly higher than A = -51.72 dBm in [RMW12]. If the model path loss constant P = 3, as in `zbfind` revision r47, maximum estimated distance increases from 20.4 m to 22.6 m. Even the updated maximum indoor distance estimate is shorter than distances examined herein, therefore the log-distance path loss model requires a smaller value for P to fit experiment results.

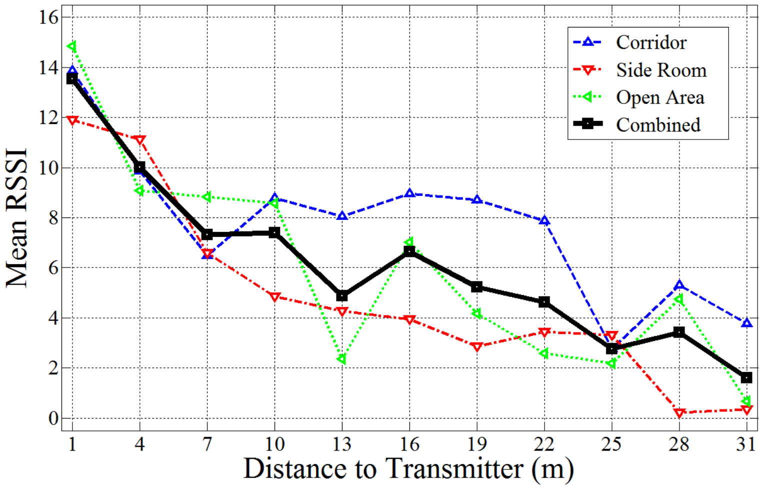


Figure 5.7: Mean RSSI versus distance measurements in a military hospital [RMLS14].

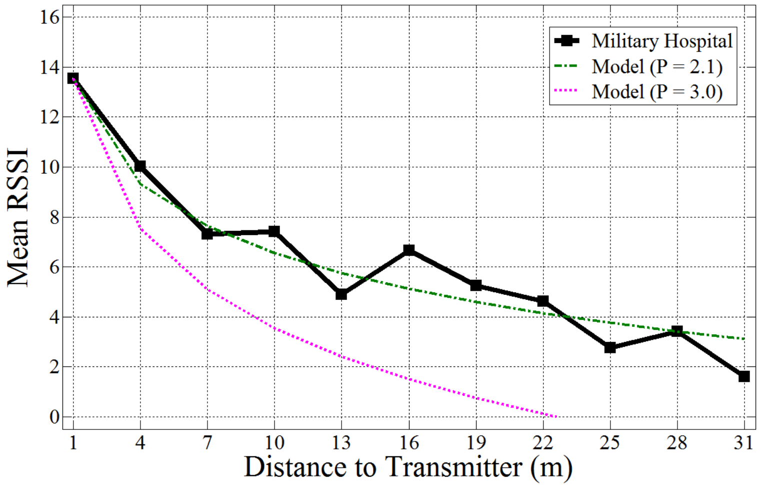


Figure 5.8: Observed RSSI and model predictions in a military hospital [RMLS14].

Fig. 5.8 overlays the combined mean RSSI measurements in Fig. 5.7 with predictions from two distance estimation models. One model uses path loss constant $P = 2.1$ (best fit to the military hospital data) and the other uses $P = 3.0$ (as in `zbfind` r47). The tuned log distance path loss model fits the real-world RSSI measurements with significantly lower

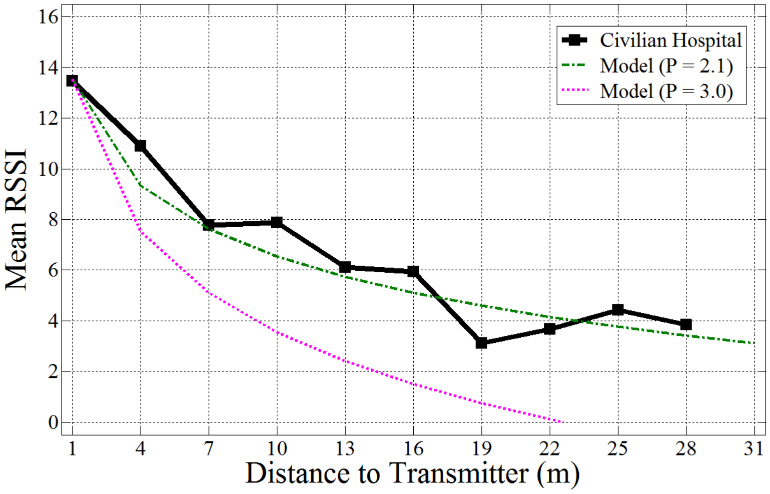


Figure 5.9: Observed RSSI and model predictions in a civilian hospital [RMLS14].

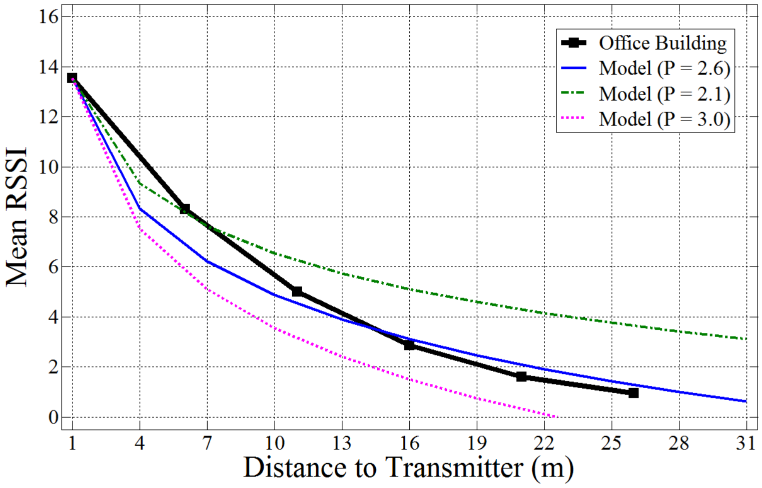


Figure 5.10: RSSI and model predictions in office buildings [RMLS14].

error than the `zbfind` r47 model. Fig. 5.9 overlays the same two distance estimation models over the civilian hospital data. Here, too, the $P = 2.1$ model appears to reduce distance estimation error over the $P = 3.0$ model. Table 5.1 quantifies distance estimation error as MAPE for the military and civilian hospital scenarios. The $P = 2.1$ model reduces

Table 5.1: Distance estimation error in hospitals.

P	Distance	Military Hospital	Civilian Hospital
2.1	$d \leq 16$ m	MAPE = 23.5%	MAPE = 25.8%
2.1	$d \leq 28$ m	MAPE = 23.1%	MAPE = 54.9%
3.0	$d \leq 16$ m	MAPE = 43.2%	MAPE = 54.2%
3.0	$d \leq 28$ m	MAPE = 50.2%	MAPE = 53.8%

error from $M = 43.2\%$ to $M = 23.5\%$ in the military hospital and from from $M = 54.2\%$ to $M = 25.8\%$ in the civilian hospital for $d \leq 16$ m. Furthermore, error reduces from 50.2% to 23.1% in the military hospital for $d \leq 28$ m, but long-range error in the civilian hospital varies by only 1.1% between the two models.

Hospitals feature wide corridors and open waiting areas that do not appear as frequently in office buildings. The path loss constant P may have a higher value in office buildings, where corridors are narrower and there is more compartmentalization of space than in hospitals. Results in Fig. 5.10 are consistent with this hypothesis, illustrating the best fit reported in [RMW12] for office building warwalking of $P = 2.6$. The model curves for $P = 2.1$ and $P = 3.0$ appear to overestimate and underestimate distance, respectively. Given these findings, use of $P = 2.1$ for hospital warwalking and $P = 2.6$ are recommended inside office buildings for the most accurate distance estimates currently available.

5.5 Smart Meter Distance Estimation

ZigBee-enabled smart meters operate from the outdoor walls of homes and businesses, yet must also be able to communicate wirelessly with any smart appliances within their Home Area Network. Smart meters address this by transmitting at higher

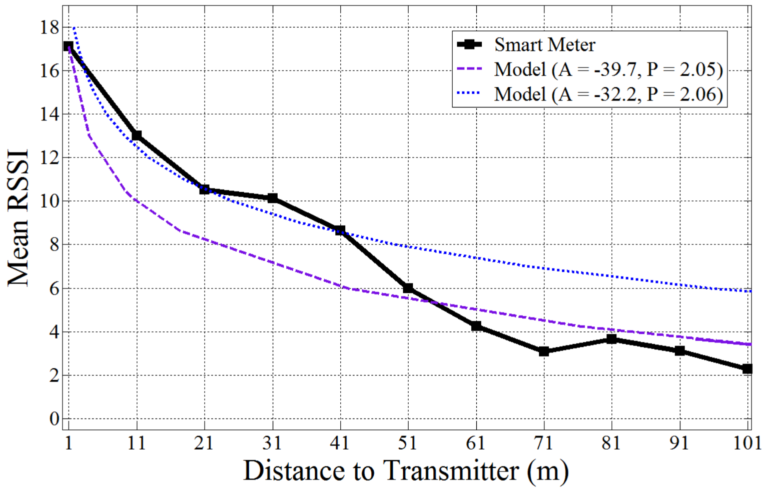


Figure 5.11: RSSI and model predictions for smart meters [RMLS14].

power than the 1 mW indoor devices considered in Section 5.4. A study involving two hundred thousand smart meters reports a median transmit 2.4 GHz ZigBee power of 66.1 mW [Ins10]. This significantly higher transmit power necessitates different values for reference RSSI A and path loss constant P in the outdoor distance estimation model.

RSSI measurements are recorded at ten-meter warwalking increments d = 1 m to d = 100 m from an Itron model CP2SOA smart electric meter. The meter is mounted outdoors on the side of an industrial warehouse. Twenty-five RSSI measurements are recorded at each distance.

Fig. 5.11 presents mean RSSI versus distance measurements for the Itron CP2SOA transmitter. If mean RSSI at d = 1 m serves as A, the best path loss constant fit is A =-39.7 dBm and P = 2.05. However, if both parameters are tuned, the best fit model utilizes A = -32.2 dBm and P = 2.06. The path loss constants in these two models are reasonable at slightly higher than free space (P = 2.0), given the presence of trees, cars, and other obstructions that impede line-of-sight. Selection of A has a significant influence on distance estimation error for $d \leq 101$ m; MAPE M is 47.4% for the A = -32.2 dBm model, versus M = 74.9% for the A = -39.7 dBm model. Additional evidence in favor of the alternative A = -32.2 dBm model is that this reference RSSI is higher, matching expected signal strength from a 66.1 mW transmitter. Given these results, model parameters A = -32.2 dBm and P = 2.06 are recommended for `zbfind` warwalking against smart utility meters.

5.6 Conclusion

This work is the first to investigate the effectiveness of CC2420-based transceiver boards for use with the `zbfind` warwalking tool. The data strongly suggest that the only KillerBee-supported hardware currently viable for use with `zbfind` is the Atmel RZUSBstick. Best-fit parameters for the log-distance path loss model are established from RSSI data collected during warwalks. Results demonstrate that the updated model parameters significantly improve distance estimation.

VI. Conclusion

Wireless security based solely on encryption keys, access control lists, and other bit-layer defenses can be made more robust by leveraging aspects of the physical layer. Such exploitation provides security measures that are simply not possible from a bit-only perspective. PHY manipulation described in Chapter 2 can be leveraged to identify the *true* transceiver type within a remote device, even if the remote device presents a spoofed Organizational Unique Identifier in its claimed MAC address. RF fingerprinting described in Chapters 3 and 4 can identify the *true* source of a spoofed transmission, even among devices from the same manufacturer. Furthermore, physical security audits and intruder localization are both made significantly easier with accurate RSSI-based distance estimation models, such as those developed in Chapter 5.

6.1 Research Contributions

The potential for novel physical layer security techniques remains largely unexplored. Processes investigated in this dissertation have spurred a series of scientific publications.

6.1.1 PHY Manipulation.

A practical demonstration of PHY manipulation was first demonstrated in [RM13], with an emphasis on the technique's potential for obscuring sensitive data (e.g., encryption keys) from eavesdroppers. Results in Chapter 2 demonstrate that true IEEE 802.15.4 transciever type can be established among six device classes with greater than 99% accuracy. A follow-up publication demonstrated that PHY manipulation is also effective for IEEE 802.11b transceivers [KRM14b]; a revised version is to appear in the Journal of Information Warfare [KRM14a]. These works confirm the hypothesis that PHY manipulation is an effective technique for physical layer exploitation. More advanced

PHY manipulation techniques also warrant investigation, such as subtle deviations from the IEEE 802.15.4 symbol-to-chip mapping sequence.

6.1.2 RF Fingerprinting.

Preliminary adaptation of RF-DNA fingerprinting techniques to IEEE 802.15.4 PHY responses for device classification was presented in [RTM12]. The addition of device verification appeared in [DRT12]. Verification resilience in response to indoor device mobility was demonstrated in [DRT13]. Investigation of decision tree classifiers for IEEE 802.15.4 fingerprinting appeared in [PTBR14a] and a revised version is to appear in the Journal of Information Warfare [PTBR14b]. Random forest techniques for differentiating LR-WPAN transmitter RF fingerprints collected for Chapter 4 will be presented at the 2014 Military Communications Conference [PTR14]. These works confirm that RF fingerprinting with greater than 90% accuracy is possible, even when a low-cost signal receiver is used.

6.1.3 RSSI and KillerBee.

Exploratory analysis toward improving the log-distance path loss model in `zbfind` was presented in [RMW12]. Observation of real-world IEEE 802.15.4 network traffic using KillerBee in conjunction with additional hardware platforms was published in [RMSB13]. Chapter 5 extends both earlier works by improving `zbfind` distance estimation by 20% and will be presented at the 2014 Military Communications Conference [RMLS14]. These works confirm that RSSI-based exploitation can be an effective tool for improved cyberspace situational awareness.

6.2 Recommendations for Future Work

6.2.1 PHY Manipulation.

Results in Chapter 2 establish that PHY manipulation techniques are likely to be effective for a diverse range of wireless network protocols. Near-term work will continue to investigate the applicability of PHY manipulation to other Local Area Network protocols, such as IEEE 802.11g, IEEE 802.11a, and IEEE 802.11ac. Similarly, cellular telephone protocols utilize PHY preambles that may prove useful for differentiating among transceiver variants.

6.2.2 RF Fingerprinting.

A significant benchmark for upcoming RF fingerprinting research is the establishment of a real-time network protection system that can accept or reject wireless traffic as valid or spoofed. Such work is already underway, and will doubtlessly leverage and improve upon the techniques reported in Chapter 3 and Chapter 4. Future work should also investigate the potential for synergy when PHY manipulation and RF fingerprinting techniques are combined. For example, RF fingerprinting accuracy should improve significantly if every device in the LR-WPAN utilizes a minor and device-unique change to the PHY preamble of outgoing transmissions that does not impact communication. Table 2.4 indicates that up to four of the PHY preamble nibbles can deviate from the standard without impacting packet reception on CC2420 devices. Preliminary work toward the quantification of this benefit, particularly at low signal to noise ratios, is already underway.

6.2.3 RSSI and KillerBee.

Results in Chapter 5 conclusively demonstrate that current CC2420-based transceiver boards are not effective for RSSI-based rangefinding at warwalking distances. However, there are other hardware platforms that can be incorporated into KillerBee in the future and evaluated for use with *zbfind*. Proposed improvements to *zbfind* in Chapter 5 will be incorporated into KillerBee upon publication.

There is also a great deal of work to be done to further improve the bit-layer effectiveness of KillerBee. The author contributed three significant improvements to KillerBee source code to date: revisions r33 (May 2012), r47 (September 2012), and r48 (October 2012). KillerBee was originally designed with an emphasis on the ZigBee protocol stack, however newer IEEE 802.15.4-based protocols including WirelessHART and ISA 100.11a are increasingly prevalent in industrial control systems. Advanced tools should be created within the KillerBee framework to specifically target and explore these newer protocols. As Wright's Law states [GBM+12]: "Practical security does not improve until tools for exploration of the attack surface are made available."

Bibliography

[AHP13] L. Apa, C. Hollman, and Mario P. Compromising industrial facilities from 40 miles away. *IOActive Technical White Paper*, 2013. https://media.blackhat.com/us-13/US-13-Apa-Compromising-Industrial-Facilities-From-40-Miles-Away-WP.pdf.

[Atm09] Atmel. *AVR Low Power 2.4 GHz Transceiver for ZigBee, IEEE 802.15.4, 6LoWPAN, RF4CE and ISM Applications*, Apr 2009. http://www.atmel.com/images/doc8111.pdf.

[BBGO08] V. Brik, S. Banerjee, M. Gruteser, and S. Oh. Wireless device identification with radiometric signatures. In *Proceedings of the 14th ACM International Conference on Mobile Computing and Networking*, pages 116–127, 2008.

[BD13] R. Booton and C. Dickens. Method of estimating satellite link quality in a time slotted tactical uhf satcom system. In *Military Communications Conference (MILCOM)*, pages 628–633, Nov 2013.

[CPMM11] M. Conti, R. Di Pietro, L. Mancini, and A. Mei. Distributed detection of clone attacks in wireless sensor networks. *IEEE Transactions on Dependable and Secure Computing*, 8(5):685–698, Sep 2011.

[DC09] B. Danev and S. Capkun. Transient-based identification of wireless sensor nodes. In *International Conference on Information Processing in Sensor Networks (IPSN)*, pages 25–36, Apr 2009.

[DDT11] R. Daidone, G. Dini, and M. Tiloca. On experimentally evaluating the impact of security on ieee 802.15.4 networks. In *Proceedings of the International Conference on Distributed Computing in Sensor Systems (DCOSS)*, pages 1–6, Jun 2011.

[DHS99] R. Duda, P. Hart, and D. Stork. *Pattern classification*. John Wiley & Sons, 1999.

[DLCED10] B. Danev, H. Luecken, S. Capkun, and K. El Defrawy. Attacks on physical-layer identification. In *Proceedings of the Third ACM Conference on Wireless Network Security*, WiSec, pages 89–98, New York, NY, USA, 2010.

[DRT12] C. Dubendorfer, B. Ramsey, and M. Temple. An rf-dna verification process for zigbee networks. In *Military Communications Conference (MILCOM)*, pages 1–6, Oct 2012.

[DRT13] C. Dubendorfer, B. Ramsey, and M. Temple. Zigbee device verification for securing industrial control and building automation systems. In *Critical Infrastructure Protection VII*, pages 47–62. Springer, 2013.

[Fre08] Freescale. *Freescale BeeStack Application Development Guide*, Jan 2008.

[GBM+11] T. Goodspeed, S. Bratus, R. Melgares, R. Shapiro, and R. Speers. Packets in packets: Orson welles' in-band signaling attacks for modern radios. In *5th USENIX Workshop on Offensive Technologies (WOOT)*, pages 54–61, 2011.

[GBM+12] T. Goodspeed, S. Bratus, R. Melgares, R. Speers, and S. Smith. Api-do: Tools for exploring the wireless attack surface in smart meters. In *45th Hawaii International Conference on System Science (HICSS)*, pages 2133–2140, Jan 2012.

[Geo10] L. George. Tracking medical equipment using radio waves, Oct 2010.

[Goo09] T. Goodspeed. Extracting keys from second generation zigbee chips. *Black Hat USA*, 2009.

[GWM+02] Q. Guo, W. Wu, D. Massart, C. Boucon, and S. De Jong. Feature selection in principal component analysis of analytical data. *Chemometrics and Intelligent Laboratory Systems*, 61(1):123–132, 2002.

[HFT09] T. Hastie, J. Friedman, and R. Tibshirani. *The elements of statistical learning*, volume 2. Springer, 2009.

[HLK+10] J. Han, J. Lee, T. Kwon, D. Jo, T. Ha, and Y. Choi. How to mitigate signal dragging during wardriving. *IEEE Pervasive Computing*, 9(1):20–27, 2010.

[HWT11] P. Harmer, M. Williams, and M. Temple. Using de-optimized lfs processing to enhance 4g communication security. In *Proceedings of 20th International Conference on Computer Communications and Networks (ICCCN)*, pages 1–8, Jul 2011.

[HY12] N. Hu and Y. Yao. Identification of legacy radios in a cognitive radio network using a radio frequency fingerprinting based method. In *IEEE International Conference on Communications (ICC)*, pages 1597–1602, 2012.

[Ins09] Texas Instruments. *CC1000 Single Chip Very Low Power RF Transceiver*, 2009. http://www.ti.com/lit/ds/symlink/cc1000.pdf.

[Ins10] Electric Power Research Institute. An investigation of radiofrequency fields associated with the itron smart meter, Dec 2010. http://smartgridcc.org/wp-content/uploads/2012/08/000000000001021126.pdf.

[Ins12] Texas Instruments. *RemoTI Developer's Guide*, 2012. http://www.ti.com/lit/ml/swru198d/swru198d.pdf.

[Ins14] Texas Instruments. *2.4 GHz IEEE 802.15.4 / ZigBee-ready RF Transceiver*, 2014. http://www.ti.com/lit/ds/symlink/cc2420.pdf.

[KK10] D. Knox and T. Kunz. Agc-based rf fingerprints in wireless sensor networks for authentication. In *IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks (WoWMoM)*, pages 1–6, Jun 2010.

[KP12] C. Kiraly and G. Picco. Where's the mote? ask the motehunter! In *IEEE Conference on Local Computer Networks Workshops (LCN Workshops)*, pages 982–990, Oct 2012.

[KRM14a] N. Kulesza, B. Ramsey, and B. Mullins. Radio frequency fingerprinting through preamble manipulation. *Journal of Information Warfare*, 2014. *To Appear*.

[KRM14b] N. Kulesza, B. Ramsey, and B. Mullins. Wireless intrusion detection through preamble manipulation. In *9th International Conference on Cyber Warfare & Security*, pages 132–139. Academic Conferences Limited, 2014.

[KS99] D. Kaplan and D. Stanhope. Waveform collection for use in wireless telephone identification, Dec 1999. US Patent 5,999,806.

[KS05] S. Kalia and M. Singh. Masking approach to secure systems from operating system fingerprinting. In *IEEE Region 10 TENCON*, pages 1–6, Nov 2005.

[KTMR09] R. Klein, M. Temple, M. Mendenhall, and D. Reising. Sensitivity analysis of burst detection and rf fingerprinting classification performance. In *IEEE International Conference on Communications (ICC)*, pages 1–5, Jun 2009.

[Kul14] N. Kulesza. Radio frequency fingerprinting techniques through preamble modification in ieee 802.11b. Masters thesis, Air Force Institute of Technology, Jun 2014.

[LGZC13] Q. Li, W. Gao, S. Zhu, and G. Cao. To lie or to comply: Defending against flood attacks in disruption tolerant networks. *IEEE Transactions on Dependable and Secure Computing*, 10(3):168–182, May 2013.

[MKP^{+}12] R. Martin, A. King, J. Pennington, R. Thomas, R. Lenahan, and C. Lawyer. Modeling and mitigating noise and nuisance parameters in received signal strength positioning. *IEEE Transactions on Signal Processing*, 60(10):5451–5463, 2012.

[MLLP12] B. Muntwyler, V. Lenders, F. Legendre, and B. Plattner. Obfuscating ieee 802.15.4 communication using secret spreading codes. In *9th Annual*

Conference on Wireless On-demand Network Systems and Services (WONS), pages 1–8, Jan 2012.

[NIS10] NIST. Guidelines for smart grid cyber security: Vol. 3, supportive analyses and references, Aug 2010.

[NW04] S. Naveen and D. Wagner. Security considerations for ieee 802.15.4 networks. In *Proceedings of the 3rd ACM Workshop on Wireless Security*, WiSec '04, pages 32–42, New York, NY, USA, 2004.

[PDG11] A. Polak, S. Dolatshahi, and D. Goeckel. Identifying wireless users via transmitter imperfections. *IEEE Journal on Selected Areas in Communications*, 29(7):1469–1479, 2011.

[PTBR14a] H. Patel, M. Temple, R. Baldwin, and B. Ramsey. Application of ensemble decision tree classifiers to zigbee device network authentication using rf-dna fingerprinting. In *9th International Conference on Cyber Warfare & Security (ICCWS)*, pages 176–186. Academic Conferences Limited, 2014.

[PTBR14b] H. Patel, M. Temple, R. Baldwin, and B. Ramsey. Introduction of random forest classifier to zigbee device network authentication using rf-dna fingerprinting. *Journal of Information Warfare*, 2014. *To Appear*.

[PTR14] H. Patel, M. Temple, and B. Ramsey. Comparison of high-end and low-end receivers for rf-dna fingerprinting. In *Military Communications Conference (MILCOM)*, Oct 2014. To Appear.

[RC07] K. Rasmussen and S. Capkun. Implications of radio fingerprinting on the security of sensor networks. In *3rd International Conference on Security and Privacy in Communications Networks (SecureComm)*, pages 331–340, Sep 2007.

[RM13] B. Ramsey and B. Mullins. Defensive rekeying strategies for physical-layer-monitored low-rate wireless personal area networks. In *Critical Infrastructure Protection VII*, pages 63–79. Springer, 2013.

[RMLS14] B. Ramsey, B. Mullins, W. Lowder, and R. Speers. Sharpening the stinger: Tuning killerbee for critical infrastructure warwalking. In *Military Communications Conference (MILCOM)*, Oct 2014. To Appear.

[RMSB13] B. Ramsey, B. Mullins, R. Speers, and K. Batterton. Watching for weakness in wild wpans. In *Military Communications Conference (MILCOM)*, pages 1404–1409. IEEE, Nov 2013.

[RMW12] B. Ramsey, B. Mullins, and E. White. Improved tools for indoor zigbee warwalking. In *37th IEEE Conference on Local Computer Networks (LCN Workshops)*, pages 921–924, Oct 2012.

[RSC12] S. Rehman, K. Sowerby, and C. Coghill. Analysis of receiver front end on the performance of rf fingerprinting. In *IEEE 23rd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, pages 2494–2499, 2012.

[RSC14] S. Rehman, K. Sowerby, and C. Coghill. Analysis of impersonation attacks on systems using rf fingerprinting and low-end receivers. *Journal of Computer and System Sciences*, 80(3):591–601, 2014.

[RTM12] B. Ramsey, M. Temple, and B. Mullins. Phy foundation for multi-factor zigbee node authentication. In *IEEE Global Communications Conference (GLOBECOM)*, pages 795–800, Dec 2012.

[RTO12] D. Reising, M. Temple, and M. Oxley. Gabor-based rf-dna fingerprinting for classifying 802.16e wimax mobile subscribers. In *International Conference on Computing, Networking and Communications (ICNC)*, pages 7–13, Jan 2012.

[SCTD09] T. Schmid, L. Choong, M. Tadjikov, and K. Dabcevic. Decoder of ieee 802.15.4 radio packets, 2009. Accessed: 2014-06-23.

[Soc06] IEEE Computer Society. Ieee standard for information technology– local and metropolitan area networks– specific requirements– part 15.4: Wireless medium access control (mac) and physical layer (phy) specifications for low rate wireless personal area networks (wpans). *IEEE Std 802.15.4-2006 (Revision of IEEE Std 802.15.4-2003)*, pages 1–320, Sep 2006.

[SSS07] T. Schmid, O. Sekkat, and M. Srivastava. An experimental study of network performance impact of increased latency in software defined radios. In *Proceedings of the Second ACM International Workshop on Wireless Network Testbeds, Experimental Evaluation and Characterization*, WinTECH '07, pages 59–66, New York, NY, USA, 2007.

[STMM08] W. Suski, M. Temple, M. Mendenhall, and R. Mills. Using spectral fingerprints to improve wireless network security. In *IEEE Global Communications Conference (GLOBECOM)*, pages 1–5, Nov 2008.

[TC13] V. Targon and A. Cavallaro. Distributed measurement selection for energy-efficient radio tracking. In *16th International Conference on Information Fusion (FUSION)*, pages 714–721, 2013.

[Tec09] Agilent Technologies. *Agilent E3238S/35688E Signal Intercept and Collection System*, 2009. http://cp.literature.agilent.com/litweb/pdf/5989-1505EN.pdf.

[Tha12] R. Thandee. Ieee 802.15. 4 implementation on an embedded device. Masters thesis, Virginia Polytechnic Institute and State University, Apr 2012.

[WLY+14] X. Wang, Y. Liu, Z. Yang, K. Lu, and J. Luo. Robust component-based localizationin sparse networks. *IEEE Transactions on Parallel and Distributed Systems*, 25(5):1317–1327, 2014.

[Wri09] J. Wright. Killerbee: practical zigbee exploitation framework. In *11th ToorCon conference, San Diego*, 2009.

[YCTC13] J. Yang, Y. Chen, W. Trappe, and J. Cheng. Detection and localization of multiple spoofing attackers in wireless networks. *IEEE Transactions on Parallel and Distributed Systems*, 24(1):44–58, Jan 2013.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704–0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704–0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202–4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD–MM–YYYY)* 18–09–2014	**2. REPORT TYPE** Doctoral Dissertation	**3. DATES COVERED** *(From — To)* Aug 2011–Aug 2014

4. TITLE AND SUBTITLE Improved Wireless Security through Physical Layer Protocol Manipulation and Radio Frequency Fingerprinting	**5a. CONTRACT NUMBER** **5b. GRANT NUMBER** **5c. PROGRAM ELEMENT NUMBER**
6. AUTHOR(S) Ramsey, Benjamin W., Captain, USAF	**5d. PROJECT NUMBER** 14G216C **5e. TASK NUMBER** **5f. WORK UNIT NUMBER**
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way WPAFB, OH 45433-7765	**8. PERFORMING ORGANIZATION REPORT NUMBER** AFIT-ENG-DS-14-S-10
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Department of Homeland Security ICS-CERT POC: Neil Hershfield, DHS ICS-CERT Technical Lead ATTN: NPPD/CS&C/NCSD/US-CERT Mailstop: 0635, 245 Murray Lane, SW, Bldg 410, Washington, DC 20528 Email: isc-cert@dhs.gov Phone: 1-877-776-7585	**10. SPONSOR/MONITOR'S ACRONYM(S)** DHS ICS-CERT **11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

12. DISTRIBUTION / AVAILABILITY STATEMENT
DISTRIBUTION STATEMENT A:
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

13. SUPPLEMENTARY NOTES
This work is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

14. ABSTRACT

Wireless networks are particularly vulnerable to spoofing and route poisoning attacks due to the contested transmission medium. Traditional bit-layer defenses including encryption keys and MAC address control lists are vulnerable to extraction and identity spoofing, respectively. This dissertation explores three novel strategies to leverage the wireless physical layer to improve security in low-rate wireless personal area networks. The first, physical layer protocol manipulation, identifies true transceiver design within remote devices through analysis of replies in response to packets transmitted with modified physical layer headers. Results herein demonstrate a methodology that correctly differentiates among six IEEE 802.15.4 transceiver classes with greater than 99% accuracy, regardless of claimed bit-layer identity. The second strategy, radio frequency fingerprinting, accurately identifies the true source of every wireless transmission in a network, even among devices of the same design and manufacturer. Results suggest that even low-cost signal collection receivers can achieve greater than 90% authentication accuracy within a defense system based on radio frequency fingerprinting. The third strategy, based on received signal strength quantification, can be leveraged to rapidly locate suspicious transmission sources and to perform physical security audits of critical networks. Results herein reduce mean absolute percentage error of a widely-utilized distance estimation model 20% by examining signal strength measurements from real-world networks in a military hospital and a civilian hospital.

15. SUBJECT TERMS
Wireless Networks, Security, Physical Layer, LR-WPAN, ZigBee

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Dr. Barry E. Mullins (ENG)
U	U	U	UU	127	**19b. TELEPHONE NUMBER** *(include area code)* (937) 255-3636 x7979 Barry.Mullins@afit.edu